

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:37:06 ; Search time 17 Seconds  
(without alignments)  
1363.013 Million cell updates/sec

Title: US-10-063-686-32

Perfect score: 445  
Sequence: 1 MSGRPTLGLALALSLA.....QDDKSLNKESTELQALVR 445

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size: 6

Total number of hits satisfying chosen parameters: 1355

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	1.8	147	1 LYC3 SHEEP	O9TUN1 ovine aries
2	8	1.8	268	1 TRPA_HABIN	P43759 haemophilus
3	8	1.8	468	1 BMP3_RAT	P49002 rattus norv
4	8	1.8	472	1 BMP3_HUMAN	P12645 homo sapien
5	8	1.8	962	1 MML9_MYCTU	P95235 mycobacteri
6	7	1.6	89	1 YHUA_BACSU	O07555 bacillus su
7	7	1.6	113	1 PURU_SYNPP	O55041 synchococc
8	7	1.6	132	1 PRDC_PASMU	O94558 pasteurella
9	7	1.6	138	1 PAZQ_TRIFL	O81900 citreusuri
10	7	1.6	151	1 DKS4_BUCBP	O89473 buchnera ap
11	7	1.6	161	1 LK52_LYCES	P13447 lycopersico
12	7	1.6	166	1 RL19_DEIRA	O91wb4 delnoccoc
13	7	1.6	185	1 FAPR_STRAM	O99u02 staphylococ
14	7	1.6	207	1 FAPR_STRAE	O84922 staphylococ
15	7	1.6	236	1 APLA_MOMO	P47456 mycoplasma
16	7	1.6	247	1 COBS_SALTI	O59544 mycoplasma
17	7	1.6	247	1 COBS_SALTI	O825n8 salmonella
18	7	1.6	247	1 COBS_SALTI	O05602 salmonella
19	7	1.6	234	1 HUPK_RHCCA	O30797 rhodocacter
20	7	1.6	297	1 HTPX_SRRCG	O30795 streptococc
21	7	1.6	298	1 Y025_MYCGE	P47271 mycoplasma
22	7	1.6	302	1 GLUT_ECOLI	P37902 escherichia
23	7	1.6	303	1 EST_ACTIM	P18773 acinetobact
24	7	1.6	337	1 GOGS_YEAST	P40107 saccharomyc
25	7	1.6	341	1 BET3_CANAL	O93803 candida alb
26	7	1.6	343	1 QREA_BORBU	O51053 borrelia bu
27	7	1.6	350	1 TME2_HUMAN	O99zb4 homo sapien
28	7	1.6	353	1 GBO_DROME	P32655 drosophila
29	7	1.6	377	1 RIM2_YEAST	P38127 saccharomyc
30	7	1.6	385	1 OP23_HAEIN	P46026 haemophilus
31	7	1.6	386	1 OP23_HAEIN	P46026 haemophilus
32	7	1.6	391	1 Y450_BUCAP	O8k999 buchnera ap

## ALIGNMENTS

RESULT 1	LYC3 SHEEP	STANDARD:	PRT:	147 AA.
AC	O9TUN1	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DE	28-FEB-2003 (Rel. 41, Last annotation update)			
OS	Lysozyme C 3 precursor (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase C).			
OC	Ovis aries (sheep).			
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Buteleostomi;			
CC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
CC	Bovidae; Caprinae; Ovis.			
CC	NCBI_TaxID=9940;			
CC	[1]			
CC	SEQUENCE FROM N.A.			
CC	MEDLINE=2008697; PubMed=10620405;			
CC	Wen Y, Irwin D.M.;			
CC	"Mosaic evolution of ruminant stomach lysozyme genes.";			
CC	Mol. Phylogenet. Evol. 13:474-482 (1999).			
CC	-1- FUNCTION: Lysozymes have primarily a bacteriolytic function; those			
CC	in tissues and body fluids are associated with the monocyte-			
CC	macrophage system and enhance the activity of immunogens.			
CC	-1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-			
CC	acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan			
CC	heteropolymers of the prokaryotes cell walls.			
CC	-1- SUBUNIT: Monomer.			
CC	-1- TISSUE SPECIFICITY: Expressed in stomach.			
CC	-1- MISCELLANEOUS: Lysozyme C is capable of both hydrolysis and			
CC	transglycosylation; it shows also a slight esterase activity. It			
CC	acts rapidly on both peptide-substituted and unsubstituted			
CC	peptidoglycan, and slowly on chitin oligosaccharides.			
CC	-1- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.			
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; AF170559; AAD51637.1; -			
CC	EMBL; AF170556; AAD51637.1; JOINED.			
CC	EMBL; AF170557; AAD51637.1; JOINED.			
CC	EMBL; AF170558; AAD51637.1; JOINED.			
CC	HSSP; P00695; 1L26.			
CC	InterPro; IPR001916; Glyco_hydro_22.			
CC	Pfam; PF00062; 1ys; 1.			
CC	PRINTS; PR00135; LY2LAC.			
CC	SMART; SMO0263; LY2L; 1.			
CC	PROSITE; PS00128; LACTALBUMIN LYSOZYME; 1.			
CC	Hydrolase; Glycosidase; Bacteriolytic enzyme; Digestion; Signal;			
CC	MultiGene family.			
CC	SIGNAL			
CC	1			
CC	18			
CC	BY SIMILARITY.			
CC	LYSOZYME C 3.			
CC	CHAIN			
CC	19			
CC	147			
CC	08k999 buchnera ap			
CC	P08001 sus scrofa			
CC	O29079 sus scrofa			
CC	P48038 oryctolagus			
CC	O89017 mus musculu			
CC	O97018 rattus norv			
CC	O53552 pseudomona			
CC	O92816 chlamydia p			
CC	O00220 homo sapien			
CC	O83361 treponema p			
CC	P57019 bacterioph			
CC	O67241 aquifex aeo			
CC	O37705 artemia san			

FT DISULFID 48 133 BY SIMILARITY.  
 FT DISULFID 83 99 BY SIMILARITY.  
 FT DISULFID 95 113 BY SIMILARITY.  
 FT ACT SITE 53 53 BY SIMILARITY.  
 FT ACT SITE 71 71 BY SIMILARITY.  
 SQ SEQUENCE 147 AA; 16265 MW; 09601329F793EBD9 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 LVIGLUF 46  
 Db 4 LVIGLUF 11

## RESULT 2

TRPA\_HABIN STANDARD; PRT; 268 AA.

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Tryptophan synthase alpha chain (EC 4.2.1.20).

GN TRPA OR H1432.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellales; Haemophilus.

NCBI\_TaxID=727;

RP SEQUENCE FROM N.A.

RC STRAIN=rd / KM20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Kiehlavsky A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Goetzue J.D.,

RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Corton M.D.,

RA Uetrechack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

RA Green C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.,

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae

RL Science 269:496-512(1995).

CC -1- FUNCTION: The alpha subunit is responsible for the aldol cleavage

CC of indoleglycerol phosphate to indole and glyceraldehyde 3-

CC phosphate.

CC -1- CATALYTIC ACTIVITY: L-serine + 1-(indol-3-yl)glycerol 3-phosphate

CC = L-tryptophan + glyceraldehyde 3-phosphate.

CC -1- PATHWAY: Tryptophan biosynthesis; fifth (last) step.

CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By

CC similarity).

CC -1- SIMILARITY: Belongs to the trpa family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U38222; AAC23079.1; -.

DR PIR; A64123; A64123.

DR HSSP; P00929; 2MSY.

DR TIGR; H11432; -.

DR HAMAP; MF\_00131; -; 1.

DR InterPro; IPR003009; FMN enzyme.

DR InterPro; IPR002028; TTP\_synthaseA.

DR Pfam; PF00290; trp\_synA; 1.

DR ProDom; PD001535; trp\_synthaseA; 1.

DR PROSITE; PS00167; TRP\_SYNTHASE\_ALPHA; 1.  
 KW Tryptophan biosynthesis; lyase; Complete proteome.  
 SQ SEQUENCE 268 AA; 28730 MW; E2FED0E5009E84 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 5.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 LSLGTAGA 151  
 Db 225 LSLGTAGA 232

## RESULT 3

BMP3\_RAT STANDARD; PRT; 468 AA.

AC P49002;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Bone morphogenetic protein 3 precursor (BMP-3).

GN BMP3 OR BMP-3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI\_TaxID=10116;

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Calvaria;

RA Takao M., Hino J., Kangawa K., Matsuo H.,

RA Chen D., Feng J.Q., Feng M., Harris M.A., Mahy P., Mundy G.R.,

RA Harris S.E.,

RT "Sequence and expression of bone morphogenetic protein 3 mRNA in

RT prolonged cultures of fetal rat calvarial osteoblasts and in rat

RT prostate adenocarcinoma PA.iii cells."

RL DNA Cell Biol. 14:235-239(1995).

CC -1- FUNCTION: Induces cartilage and bone formation.

CC -1- SUBUNIT: Homodimer; disulfide-linked (By similarity).

CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN TRACHEA, SPLEEN, SMALL

CC BONE MARROW.

CC -1- SIMILARITY: Belongs to the TGF-beta family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; D63860; BA009922.1; -.

DR EMBL; S77492; AAB33725.1; -.

DR HSSP; P12643; 3BMP.

DR InterPro; IPR001839; TGFb.

DR Pfam; PF00019; TGF-beta; 1.

DR ProDom; PD000357; TGFb; 1.

DR SMART; SM00204; TGFb; 1.

DR PROSITE; PS00250; TGF\_BETA\_1; 1.

KW Signal; Growth factor; Cytokine; Osteogenesis; Chondrogenesis;

KW Glycoprotein.

FT SIGNAL 1 22 POTENTIAL.

FT PROPEP 23 358 POTENTIAL.

FT CHAIN 359 468 BONE MORPHOGENETIC PROTEIN 3.

FT DISULFID 366 433 BY SIMILARITY.

FT DISULFID 395 465 BY SIMILARITY.

FT DISULFID 399 467 BY SIMILARITY.

FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 468 AA; 52675 MW; 05315D4954DC3CA1 CRC64;

Query Match  
Best Local Similarity 100.0%; Pred. No. 9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 SLSILFF 224  
DB 439 SLSILFF 446

## RESULT 4

BMP3\_HUMAN STANDARD; PRT; 472 AA.  
ID BMP3\_HUMAN  
AC P12645;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Bone morphogenetic protein 3 precursor (BMP-3) (Osteogenin) (BMP-3A).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89072730; PubMed=3201241;  
RA Wozney J.M., Rosen V., Celeste A.J., Mittlek L.M., Whitters M.J.,  
RA Kriz R.W., Hewick R.M., Wang E.A.;  
RT "Novel regulators of bone formation: molecular clones and  
RT activities";  
RL Science 242:1528-1534 (1988).  
CC - FUNCTION: Induces cartilage and bone formation.  
CC - SUBUNIT: Homodimer; disulfide-linked (By similarity).  
CC - TISSUE SPECIFICITY: EXPRESSED IN LUNG, OVARY AND SMALL INTESTINE.  
CC - SIMILARITY: Belongs to the TGF-beta family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
EMBL: M22491; AAA51836.1; -  
DR PIR: D37278; BMT03.  
DR HSSP: P12643; 3BMP.  
DR Genew; HGNC:1070; BMP3.  
DR MIM; 112263; -  
DR GO: GO:0005102; F:receptor binding; TAS.  
DR GO: GO:0007267; P:cell-cell signaling; TAS.  
DR GO: GO:0001501; P:skeletal development; TAS.  
DR InterPro: IPR001839; TGFb.  
DR Pfam; PF00019; TGF-beta; 1.  
DR ProDom; PD000357; TGFb; 1.  
DR SMART; SM00204; TGFb; 1.  
DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
KW Signal; Growth factor; Cytokine; Osteogenesis; Chondrogenesis;  
KW Glycoprotein.  
FT SIGNAL 1 22  
FT PROPEP 23 362  
FT CHAIN 363 472 BONE MORPHOGENETIC PROTEIN 3.  
FT DISULFID 399 469 BY SIMILARITY.  
FT DISULFID 403 471 BY SIMILARITY.  
FT DISULFID 435 436 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 472 AA; 53406 MW; 95C3B7ED5C9D596F CRC64;

Query Match  
Best Local Similarity 100.0%; Pred. No. 9.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 SLSILFF 224  
DB 443 SLSILFF 450

## RESULT 5

MML9\_MYCTU STANDARD; PRT; 962 AA.  
ID MML9\_MYCTU  
AC P95235;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Putative membrane protein mmp19.  
GN MML9 OR RV2339 OR MT2402 OR MTCY98.08.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Broese R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglsmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,  
RA Hornsby T., Jagels J., Krogh A., McLean N., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence";  
RL Nature 393:537-544 (1998).  
[2]  
SEQUENCE FROM N.A.  
RX STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=22206484; PubMed=12218036;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft M., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.,  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains";  
RL J. Bacteriol. 184:5479-5490 (2002).  
CC - SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC - SIMILARITY: Belongs to the mmpL family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
EMBL: Z83860; CAB6156.1; -  
DR EMBL: AB007080; AAK4696.1; -  
DR PIR: D70661; D70661.  
DR TIGR: MT2402; -  
DR TubercuList; RV2339; -  
DR InterPro: IPR004707; ActII.  
DR InterPro: IPR004869; MMP1.

```

DR Pfam: PF03176; MPM1, 2.
DR TIGRfam: TIGR00833; act11, 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45
FT TRANSMEM 213 233
FT TRANSMEM 256 276
FT TRANSMEM 302 322
FT TRANSMEM 335 355
FT TRANSMEM 383 403
FT TRANSMEM 768 788
FT TRANSMEM 796 816
FT TRANSMEM 820 840
FT TRANSMEM 867 887
FT TRANSMEM 916 936
FT TRANSMEM 937 957
SQ SEQUENCE 962 AA; 104612 MW; 9C8E33E4F44AC985 CRC64;

Query Match
Best Local Similarity 1.8%; Score 8; DB 1; Length 962;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 VLLSLGTA 149
DB 802 VLLSLGTA 809

RESULT 6
YHJA_BACSU STANDARD; PRT; 89 AA.
ID YHJA_BACSU STANDARD; PRT; 89 AA.
AC 00755;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yhjA precursor.
GN yhjA OR BSU10440.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Besieres P., Bolojin A., Borchert S.,
RA Borrias R., Bousnier L., Brans A., Braun M., Brimell S.C., Bron S.,
RA Brouillet S., Brushchi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codari J.J., Cornett I.F., Cummings N.J., Daniel N.M.,
RA Denizot F., Devignon K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Erington J., Fabret C., Ferrati E., Foulger D.,
RA Filiz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holstapel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasaahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
RA Paro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,

```

---

```

RA Sato T., Scanlan E., Schleich S., Schroeder R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vanlier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT subtilis".
RL Nature 390:249-256(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: Y14081; CAAT4463.1; -
DR EMBL: Z98109; CAB12884.1; -
DR PIR: A69833; A69833.
DR Subtilisin; BG13068; yhjA.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 27
FT CHAIN 28 89
FT SEQUENCE 89 AA; 9801 MW; 3980CF8D13EED66F CRC64;
SQ SEQUENCE 89 AA; 9801 MW; 3980CF8D13EED66F CRC64;

Query Match
Best Local Similarity 1.6%; Score 7; DB 1; Length 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 AVLLSLG 147
DB 6 AVLLSLG 12

RESULT 7
PURL_SYN7 STANDARD; PRT; 113 AA.
ID PURL_SYN7 STANDARD; PRT; 113 AA.
AC Q55041;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoribosylformylglycinamide synthase II (EC 6.3.5.3) (FGAM
DE synthase II) (Fragment).
GN PURL.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RC
RX
RA Liu Y., Tsinoemas N.F., Golden S.S., Kondo T., Johnson C.H.,
RT "Circadian expression of genes involved in the de novo purine
RT biosynthetic pathway of cyanobacteria."
RT Submitted (Aug-1995) to the EMBL/Genbank/DBJ databases.
CC -1- CARBOLIC ACTIVITY: ATP + N(2)-formyl-N(1)-(5-phospho-D-
CC ribosyl)glycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-
CC (formamido)-N(1)-(5-phospho-D-ribosyl)acetamide + L-glutamate.
CC -1- PATHWAY: De novo purine biosynthesis; fourth step.
CC -1- SUBUNIT: Heterodimer of two subunits, purp and purl.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the FGAMS family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----

```

DR EMBL; U33322; AA05111.2; -.  
DR HAMAP; MF 00420; -. 1.  
DR InterPro; IPR000728; AIR\_synth.  
DR Pfam; PF00586; AIRS; 1.  
KM Purine biosynthesis; Ligase; ATP-binding.  
FT NP\_BIND 107 >113 ATP (POTENTIAL).  
FT NON\_TER 113 113  
SQ SEQUENCE 113 AA; 12067 MW; 901AD1692AF7B8AD CRC64;

Query Match 1.6%; Score 7; DB 1; Length 113;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 AISSAPP 120  
DB 3 AISSAPP 9

RESULT 8  
FRDC\_PASMTU STANDARD; PRT; 132 AA.  
AC O9CP58;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Fumarate reductase subunit C.  
GN FRDC OR PM0199.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Pasteurella.  
OX NCBI\_TaxId=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM70;  
RA MEDLINE=21145866; PubMed=11248100;  
RX May B.J., Zhang Q., Li L.L., Paustian M.L., Whittem T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida PM70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).  
CC -1- FUNCTION: Seems to be involved in the anchoring of the catalytic  
CC components of the fumarate reductase complex to the cytoplasmic  
CC membrane (by similarity).  
CC -1- SUBUNIT: Part of an enzyme complex containing four subunits: a  
CC flavoprotein (frdA), an iron-sulfur protein (frdS), and two  
CC hydrophobic anchor proteins (frdc and frdd) (by similarity).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (by similarity).  
CC -1- SIMILARITY: Belongs to the frdc family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AE006054; AAK02283.1; -.  
DR HAMAP; MF 00708; -. 1.  
DR InterPro; IPR003510; Fumarate\_red\_C.  
DR Pfam; PF02300; Fumarate\_red\_C; 1.  
DR ProDom; PD015900; Fumarate\_red\_C; 1.  
KM Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 34 56  
FT TRANSMEM 71 93  
FT TRANSMEM 106 128  
SQ SEQUENCE 132 AA; 15204 MW; 3B18C8BD30721985 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

36 FTSTVITL 127

DB 121 FTSTVITL 127

RESULT 9  
PA2O\_TRIFL STANDARD; PRT; 138 AA.  
AC O8UIG0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Phospholipase A2 isozyme cPLA-B (A) precursor (EC 3.1.1.4)  
DE (Phosphatidylcholine 2-acylhydrolase).  
OS Trimeresurus flavoviridis (Habu).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroides;  
OC Viperidae; Crotalinae; Trimeresurus.  
OX NCBI\_TaxId=88087;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Amami-Oshima, and Kagosima; TISSUE=Venom gland;  
RX MEDLINE=22499762; PubMed=12612832;  
RA Chijiwa T., Yamaguchi Y., Ogawa T., Deshimaru M., Nobuhisa I.,  
RA Nakashima K.-I., Oda-Ueda N., Fukumaki Y., Hattori S., Ohno M.;  
RT "Intriguing evolution of Trimeresurus flavoviridis venom  
RT phospholipase A(2) isozymes.";  
RL J. Mol. Evol. 56:286-293 (2003).  
CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the  
CC 2-acyl groups in 3-sn-phosphoglycerides.  
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-  
CC acylglycerophosphocholine + a fatty acid anion.  
CC -1- COFACTOR: Binds 1 calcium ion per subunit (by similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- SIMILARITY: Belongs to the phospholipase A2 family. Group II  
CC subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB087496; BAC02719.1; -.  
DR InterPro; IPR001211; PhospholipaseA2.  
DR Pfam; PF00068; phoslip; 1.  
DR PRINTS; PR00389; PPHPLIPASEA2.  
DR ProDom; PD000303; PhospholipaseA2; 1.  
DR SMART; SM00085; PA2C; 1.  
DR PROSITE; PS00119; PA2\_ASP; 1.  
DR PROSITE; PS00118; PA2\_HIS; 1.  
KM Hydrolysis; Lipid degradation; Calcium; Multigene family; Signal.  
FT SIGNAL 1 16  
FT CHAIN 17 138  
FT ACT\_SITE 63 63  
FT ACT\_SITE 105 105  
FT DISULFID 42 131  
FT DISULFID 44 60  
FT DISULFID 59 111  
FT DISULFID 65 138  
FT DISULFID 66 104  
FT DISULFID 73 97  
FT DISULFID 91 102  
FT METAL 43 43  
FT METAL 45 45  
FT METAL 47 47  
FT METAL 49 49  
SQ SEQUENCE 138 AA; 15204 MW; 3B18C8BD30721985 CRC64;

36 FTSTVITL 127

Query Match 1.6%; Score 7; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 1.6%; Score 7; DB 1; Length 138;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 86 ITAVLV 92  
 DB 6 ITAVLV 12

## RESULT 10

DKSA\_BUCBP STANDARD; PRT; 151 AA.

AC Q89AR3; 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Dnak suppressor protein homolog.  
 GN DKSA OR BBP184.  
 OS Buchnera aphidicola (subsp. *Baijzongia pistaciace*).  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Buchnera.  
 CX NCBI\_TaxID=135842;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426901; PubMed=12522265;  
 RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Raunell C., Abascal F.,  
 RA Bactolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.T.,  
 RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.,  
 RT "Reductive genome evolution in *Buchnera aphidicola*."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).  
 CC -1- SIMILARITY: Contains 1 dksa/lrar-type zinc finger.  
 CC -1- SIMILARITY: IN THE N-TERMINAL HALF, TO THE N-TERMINAL OF  
 CC B. SUBTILIS YTEA.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE014016; AAC26916.1; -  
 DR InterPro; IPR000962; Znf\_DsKa/TRAR.  
 DR Pfam; PF01258; Zf\_dksa/TRAR; 1.  
 DR PRINTS; PR00618; DKSANZNFINGER.  
 DR PROSITE; PS01102; DKSATRAR\_ZN\_FINGER; 1.  
 KW Zinc-finger; Complete proteome.  
 FT ZN FING 114 138 TRAR/DKSA-TYPE (BY SIMILARITY).  
 SQ SEQUENCE 151 AA; 17876 MW; 77B443C75CD958FA CRC64;

Query Match 1.6%; Score 7; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 217 SSLSLTF 223  
 DB 10 SSLSLTF 16

## RESULT 11

LAS2\_LYCES STANDARD; PRT; 161 AA.

AC P13447;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Another specific LAT52 protein precursor.  
 GN LAT52

OS Lycopersicon esculentum (Tomato).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

CC lamids; Solanales; Solanaceae; Solanum.

OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. VF36;  
 RX MEDLINE=89364704; PubMed=2770694;  
 RA Twell D., Wing R., Yamaguchi J., McCormick S.;  
 RT Isolation and expression of an anther-specific gene from tomato."  
 RL Mol. Gen. Genet. 217:240-245 (1989).  
 CC -1- FUNCTION: May play a role during germination or early tube growth.  
 CC -1- TISSUE SPECIFICITY: Expressed in anthers and pollen.  
 CC -1- SIMILARITY: Belongs to the Ole e I family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X15855; CAA33854.1; -  
 DR FIR; S04765; S04765.  
 DR InterPro; IPR006041; Ole\_e1\_ext.  
 DR InterPro; IPR006040; PollenOle\_e1.  
 DR Pfam; PF01190; PollenOle\_e1; 1.  
 DR PROSITE; PS00925; OLEBI; 1.  
 KW Glycoprotein; Signal.  
 FT SIGNAL 1 17  
 FT CHAIN 18 161 POTENTIAL.  
 FT CARBOHYD 61 61 ANOTHER SPECIFIC LAT52 PROTEIN.  
 SQ SEQUENCE 161 AA; 17834 MW; 7BD5E629F69758F2 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 10 LCTIALA 16  
 DB 11 LCTIALA 17

RESULT 12  
 RL19 DEIRA STANDARD; PRT; 166 AA.  
 AC Q9RWB4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE 50S ribosomal protein L19.  
 GN RPLS OR DR0755.  
 OS Deinococcus radiodurans.  
 CC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
 CC Deinococcaceae; Deinococcus.  
 CX NCBI\_TaxID=1299;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile M., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Uretchack T., Zaleski C.,  
 RA Nakarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1."  
 RL Science 286:1571-1577(1999).  
 CC -1- FUNCTION: This protein is located at the 30S-50S ribosomal subunit  
 CC interface and may play a role in the structure and function of the  
 CC aminoacyl-tRNA binding site (by similarity).  
 CC -1- SIMILARITY: Belongs to the L19p family of ribosomal proteins.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE001931; AAF10334.1; -  
 DR PIR; G75478; G75478.  
 DR PDB; INKW; 11-FEB-03.  
 DR TIGR; DR0755; -  
 DR HAMAP; MF\_00402; -; 1.  
 DR InterPro; IPR001857; Ribosomal\_L19.  
 DR Pfam; PF01245; Ribosomal\_L19; 1.  
 DR PRINTS; PR00061; RIBOSOMAL19.  
 DR ProDom; PD002979; Ribosomal\_L19; 1.  
 DR TrEMBL; TIGR01024; rplS bact; 1.  
 DR PROSITE; PS01015; RIBOSOMAL\_L19; FALSE-NEG.  
 DR KMW Ribosomal protein; Complete proteome; 3D-structure.  
 DR KMW SEQUENCE 166 AA; 18316 MW; 1DA8E740601FFAF9 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 423 ARAQDQK 429  
 DB 129 ARAQDQK 135

RESULT 13  
 FAPR STRAM STANDARD; PRT; 185 AA.  
 ID FAPR STRAM STANDARD; PRT; 185 AA.  
 AC Q99UP0;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Transcription factor fapR (Fatty acid and phospholipid biosynthesis  
 DE regulator).  
 GN FAPR OR SAV1228 OR SA1071 OR MW111.  
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699),  
 OS Staphylococcus aureus (strain N315), and  
 OS Staphylococcus aureus (strain MW2).  
 CC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 CC NCBI\_TaxID=158878, 158879, 196620;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mu50 / ATCC 700699, and N315;  
 RC MEDLINE=21311952; PubMed=1148146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,  
 RA Karamori M., Matsunari H., Murayama A., Murakami H., Hoshiyama A.,  
 RA Mikutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.,  
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
 RT aureus".  
 RL Lancet 357:1225-1240(2001).  
 RU [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mu2;  
 RC MEDLINE=22040717; PubMed=12044378;  
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
 RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,  
 RA Yamamoto K., Hiratake K.,  
 RT "Genome and virulence determinants of high virulence community-  
 RT acquired MRSA".  
 RL Lancet 359:1819-1827(2002).  
 CC -1- FUNCTION: Transcriptional factor involved in regulation of  
 CC membrane lipid biosynthesis by repressing genes involved in fatty

CC -----  
 CC acid and phospholipid metabolism (By similarity).  
 CC -1- SIMILARITY: Belongs to the fapR family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP003361; BAB57390.1; -  
 DR EMBL; AP003133; BAB42323.1; -  
 DR EMBL; AP004826; BAB94976.1; -  
 DR PIR; G89895; G89895.  
 DR HAMAP; MF\_01814; -; 1.  
 DR KMW Fatty acid biosynthesis; Transcription regulation; Repressor;  
 DR KMW DNA-binding; Complete proteome.  
 DR KMW SEQUENCE 185 AA; 21434 MW; F779DC3C4D94E0F CRC64;

Query Match 1.6%; Score 7; DB 1; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 LRRRIKL 103  
 DB 50 LRRRIKL 56

RESULT 14  
 FAPR STRAP STANDARD; PRT; 186 AA.  
 ID FAPR STRAP STANDARD; PRT; 186 AA.  
 AC Q8CSV2;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Transcription factor fapR (Fatty acid and phospholipid biosynthesis  
 DE regulator).  
 GN FAPR OR SE0903.  
 OS Staphylococcus epidermidis.  
 CC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 CC NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 12228;  
 RC PubMed=12950922;  
 RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,  
 RA Qian Z.-Q., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.,  
 RA Yuan Z.-H., Zhao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,  
 RT "Genome-based analysis of virulence genes in a non-biofilm-forming  
 RT Staphylococcus epidermidis strain (ATCC 12228)".  
 RL Mol. Microbiol. 49:1577-1593(2003).  
 CC -1- FUNCTION: Transcriptional factor involved in regulation of  
 CC membrane lipid biosynthesis by repressing genes involved in fatty  
 CC acid and phospholipid metabolism (By similarity).  
 CC -1- SIMILARITY: Belongs to the fapR family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE016746; AA004500.1; ALT\_INIT.  
 DR HAMAP; MF\_01814; -; 1.  
 DR KMW Fatty acid biosynthesis; Transcription regulation; Repressor;  
 DR KMW DNA-binding; Complete proteome.  
 DR KMW SEQUENCE 186 AA; 21421 MW; E817002658A461AE CRC64;

CC -----  
 CC membrane lipid biosynthesis by repressing genes involved in fatty

Query Match 1.6%; Score 7; DB 1; Length 186;  
 Best Local Similarity 100.0%; Pred. No. 42;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 97 LRRIKL 103  
|||  
Db 50 LRRIKL 56

RESULT 15  
Y214\_MYCGE STANDARD; PRT; 207 AA.  
ID Y214\_MYCGE  
AC P47456;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MG214.  
GN MG214.  
OS Mycoplasma genitalium.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
CX NCBI\_TaxID=2097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33530 / G-37;  
RX MEDLINE=96026346; PubMed=7569993;  
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Fuhmann J.L.,  
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,  
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
RA Tomblin J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lueker T.S.,  
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
RT "The minimal gene complement of Mycoplasma genitalium.";  
RL Science 270:197-403(1995).  
CC -1 SIMILARITY: TO B.SUBTILIS YPUH.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U39701; AAC71433.1; -.  
DR FIK; F64223; F64223.  
DR TIGR; MG214; -.  
DR InterPro; IPR005234; Cons\_hypoth281.  
DR Pfam; PF04079; DUF387; 1.  
DR TIGRFAMs; TIGR00281; TIGR00281; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 207 AA; 23343 MW; 24741D38C57B56F CRC64;  
Query Match 1.6%; Score 7; DB 1; Length 207;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 43 GLIFVCG 49  
|||  
Db 27 GLIFVCG 33

Search completed: May 27, 2004, 16:41:00  
Job time : 19 secs







SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-702-344-20

Query Match 3.2%; Score 52.8; DB 1; Length 1389;  
Best Local Similarity 49.1%; Pred. No. 2.4e-06;  
Matches 169; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

QY 856 TCTTCTACATCAAGAACCGTTGTGAAAGGTCATTTTAACTCTGTGTGAGGATTC 915  
DB 914 TCAGGTACCAACAGAGCTCCCTGGCCCTTGGCGCCTCATCTGCGCATGTGAGATCA 973  
QY 916 CGAGATCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975  
DB 974 TCCGTGTGATCTCAAGTACCTGATCAGGCGCTGAAAGCTGACAGAA---CAAGTTG 1030  
QY 976 CCAGTACCTGTTCGAGTCTGCTACTGCTGTTTGTGCTGCTTGAACAATACCTGCTCC 1035  
DB 1031 CCAAGTGCCTCATGACCTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090  
QY 1036 ATCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1095  
DB 1091 TCCCTTAATGAGATGCTCAATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1150  
QY 1096 CAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1155  
DB 1151 CAGAGATGCTCTTCTTCCGCTCATGAGAAACATCATCAGAGTGTCTCTGATTAAG 1210  
QY 1156 TTGGAGACTTCAATTTTCTAGGAAAGCTGTGTGTGTGT 1199  
DB 1211 TTACTGACTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1254

RESULT 4  
US-08-232-463-14

Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMM  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SRO ID NO: 14;  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpt-Fls  
US-08-232-463-14

Query Match 3.0%; Score 49.8; DB 1; Length 7218;  
Best Local Similarity 2.5%; Pred. No. 7.7e-05;  
Matches 9; Conservative 213; Mismatches 145; Indels 0; Gaps 0;

QY 8 TTCTTCACTATACCCACAGTCCAAAGAGACTGCTGTCCAGGCTACCACTTC 67  
DB 1082 YY 1141  
QY 68 TCAAGCAAGTCAATTCCTTATTAACGATGTGCTCAACACAGTGAAGTCACTC 127  
DB 1142 YY 1201  
QY 128 CTAATTTGATCTGTTTGAATAAGATGTGACACCTCCACCGAATCTAAGTGAAT 167  
DB 1202 YY 1261  
QY 188 CATGTGGGAAGAGATCAATCCTGCGCTGATCTCGATTCAGCTGCTGCTGCTG 247  
DB 1262 YY 1321  
QY 248 CATGATTTTACCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 307  
DB 1322 YY 1381  
QY 308 TATTTGGATCTGTTGTTGCTGCGGCTGTTTATGCTGCTGCTGCTGCTGCTGCT 367  
DB 1382 YY 1441  
QY 368 CGACCTC 374  
DB 1442 ATTCCTC 1448

RESULT 5

US-09-621-976-2813  
Sequence 2813, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J. B.  
APPLICANT: Uopert, S.  
APPLICANT: Giordano, J. Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 2813  
LENGTH: 832  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 235..399  
US-09-621-976-2813

Query Match 2.4%; Score 39.8; DB 4; Length 832;  
Best Local Similarity 9.2%; Pred. No. 0.027;  
Matches 32; Conservative 164; Mismatches 151; Indels 0; Gaps 0;

QY 1146 ATTAACGTGTTGGAGACTTCATATTTTCTAGAGAAAGTGTAGTGTTCTACT 1205  
Db 3 WTWKRTTWYACCTWTKMSMSYMWTKYTWKRWKKAAMKTKMTWTWYRY 62  
QY 1206 GTTTTGGAGACTGATGCTTTTACTACATCGGACATCGAGTGTGGAGTCCCT 1265  
Db 63 AMWGTGKKAACRTKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 122  
QY 1266 CTTGATTTGTTAGTCTTTTGTCCCTACTAGAGCCCATAGTTTCTGTTGTTAA 1325  
Db 123 CTKKAAVYRTKTYTSSKGTWTKRWKKAAMTWTKYTWKRWKKAAMTWTKY 182  
QY 1326 ACTGTGCTGATGACATTTTCTGTTGTTGCTGTTGATCTGAAACAATAGATGATCG 1385  
Db 133 YCWWGKARKMTWTKRSYSASAKCCYSCSAGMSWKYMMWMMWMMWMMWMMWMMW 242  
QY 1386 TCGAAGAAAGCCCTACTTTATGATCAAGATTCTGAGTTTGTAAAAAGAGCAACAA 1445  
Db 243 WPAACMMRKKYAGKXKTSYKSMWMTSWMKCYTKARWGTGYCYKKGWKGKGRWYAS 302  
QY 1446 TTAACAATGCAAGGACAGACAGACAGACACATCATTAAGATGA 1492  
Db 303 KKYMKRWMMWCAWRYRYSTGRASMMWRMYTMMWMMWMMWMMWMMWMMWMMW 349

RESULT 6  
US-08-232-463-14/C  
Sequence 14, Application US/08232463  
Patent No. 5670367

GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESS: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0239

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZgpt-Fls

Query Match 2.3%; Score 38.8; DB 1; Length 7218;  
Best Local Similarity 3.4%; Pred. No. 0.28; Mismatches 0; Gaps 0;  
Matches 10; Conservative 165; Indels 117; Gaps 0;  
QY 1367 GGAACAATGATGATCGTCAGAAAAGCCCTACTTATGATCAAGATTCTGAGTTT 1426  
Db 1341 RRR 1282  
QY 1427 CTTAAAAAGAGACAAATTAAACATGCAAGGGACAGACAGACACATCATTAAG 1486  
Db 1281 RRR 1222  
QY 1487 GAATGAGAGGAGAACTCAGGCACTTGTGATGATAGATACCATTTAGTATCTGT 1546  
Db 1221 RRR 1162  
QY 1547 ACTGGAACAATTTCTCTTAAAGACCAATTTCAGATGAAATGAGACCACTAGAGA 1606  
Db 1161 RRR 1102  
QY 1607 AAGTTAGTGAATTTTCTTAAAGACCTAATTAACCTATCTCTCCCA 1658  
Db 1101 RRR 1050

RESULT 7  
US-08-658-665-53/C  
Sequence 53, Application US/08658665  
Patent No. 5997878  
GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
APPLICANT: Pincus, Steven E.  
APPLICANT: Cox, William I.  
TITLE OF INVENTION: Recombinant Poxvirus - Cytomegalovirus,  
NUMBER OF SEQUENCES: 190  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/658,665  
FILING DATE: 05-JUN-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer Bq., William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2720.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)840-0712  
TELEFAX: (212)840-3533

INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1386 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-658-665-53

Query Match 2.3%; Score 38.2; DB 2; Length 1386;  
Best Local Similarity 45.7%; Pred. No. 0.13; Mismatches 12; Gaps 0;  
Matches 12; Conservative 158; Indels 154; Gaps 0;



QY	Db	1059	AGCCTCTTCCATCTGACTCTCTCGGCGCATGGCCCGTAGGTGATTCACACTAGAGAGACA	1000
QY	639	GCTGCCCAAGTTATGAAAGCGCGCCAAATGGAATATAAGCCCTTTGGGCAVTTGGGTAC	698	
Db	999	GACTCTAGAGAGATCGGCCCCAGAAATATACCTGGGCAAGACCTTCATGACGATCTCCTC	940	
QY	699	ATGTGTCGTACACTTTATATTGGGCTACTATCTGSACTAGTAATTCATCCTTGGCTGCAG	758	
Db	939	AATCGCGGCTTCACTTACTGATACCTGAACTTCAAGGCTTGTTATGAGAGCCGCTTGGCGAG	880	
QY	759	CAAAATGACTATAGCTGGGGCAGTGGTATCTTTGTTATTTCAACAGAAAGTAAATGATCCT	818	
Db	879	CATCACACTAGTGACCTTGTACTCATTAACAATTTGTTCCACACATGTAAGAGATATAC	820	
QY	819	CCGTGATATCCCATCCTTTGGTCTCTCCCATCTCTCTTCTTCAACATCA	869	
Db	819	CATTAATATGTGATCAATGTGCGTAGACACTTATCTCTCTCTCCATCCAA	769	

## RESULT 10

US-09-916-963-53/c  
Sequence 53, Application US/09916963  
Patent No. 6632438  
GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
Pincus, Steven E.  
Kaufman, Elizabeth K.  
Cox, William I.  
TITLE OF INVENTION: RECOMBINANT POXVIRUS - CYTOMEGALOVIRUS,  
COMPOSITIONS AND USES  
NUMBER OF SEQUENCES: 176  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/916,963  
FILING DATE: 26-Jul-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,014  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer Esq., William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-0712  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1386 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 53:  
US-09-916-963-53

Accession	Sequence	Position
Dd	AGCCTCTTCTCATCTGACTCTCTCGGCGATGCGCCGATAGTCACTTCACACTTAGAGAGCA	1059
Qy	GCTGCCCAAGTTATGAGAAAGCGGCGCAAGTGAATATTAAGCCCTTTGGGCAATTGGTAC	639
Dd	GACTCTCAGAGAGATCGGGCCCCAGAAATGTACTGGGCAAGACCTTCATGCGAGATCTCTC	999
Qy	ATGTGTCTGATCCATTTAATTGCGCTCATCTGCACTAGTGAATTCATCTTCGCGTCAG	699
Dd	AATCGGGGCTTCATTAACACTGATTAACCTGAGCTTGATATCAGAGCGCGCTTGGCCAG	939
Qy	CAATGACTATAGCTGGGCGAGTGGTACTTGTATTTTCAACAGAGTAATAAATGATCCT	759
Dd	CATCAGCTAGTAGCTTTTACTCATTTACACATTTGTTCCACACATGTAAGAGATATC	879
Qy	CTGTATATCCCATCTTTTGTCTCTCTCCATCTCTTCTTTCTTCAATCAA	819
Dd	CATTAATATGATCAATATGCGCGAGACACCTTGTCTCTCTCCCATCCAA	819

## RESULT 11

```

US-08-658-665-54
Sequence 54, Application US/08658665
Patent No. 5997878
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
APPLICANT: Kauffman, Elizabeth K.
TITLE OF INVENTION: Recombinant Poxvirus - Cytomegalovirus,
TITLE OF INVENTION: Compositions and Uses
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,665
FILING DATE: 05-JUN-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer Esq., William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2720.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)840-3333
TELEFAX: (212)840-0712
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 3409 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-665-54

```

Db 818 GACTCTCAGAGATGCGCCCGCCAGAAATGACTGGCAAGACCTTCATGACATCTCTC 877  
Qy 699 ATGTGTCGTACATTTAATGCGCTCATCTGACATGATGATTCATCTTGGCGCCAG 758  
Db 878 AATGCGGCGCTTCATTAACCTGATTAACCTGAGCTTGATTAACAGAGCCGCTGGCCAG 937  
Qy 759 CAATGACTATAGCTGGGCGAGGTGTTACTGTTATTTCAACAGAAATGAATGATCCT 818  
Db 938 CATGACACTAGTACCTTGTACTCATTAACATTTTTCACACATGATGAGATATC 997  
Qy 819 CCTGATCATCCATCTTTCGTCTCTCCATCTCTTCTTACATCA 869  
Db 998 CATAAATATGATCAATGCTGAGACACTTGTCTCTCTCATCCAA 1048

RESULT 12  
US-08-796-101-18  
Sequence 18, Application US/08796101  
Patent No. 6183752

GENERAL INFORMATION:  
APPLICANT: EPSTEIN, STEPHEN E.  
APPLICANT: FINKEL, TOREN  
APPLICANT: SPIER, EDITH  
APPLICANT: ZHOU, YI FU  
APPLICANT: ZHU, JIANHUI  
APPLICANT: ERDILE, LORNE  
APPLICANT: PINCUS, STEVEN  
TITLE OF INVENTION: RESTENOSIS/ATHEROSCLEROSIS DIAGNOSIS,  
TITLE OF INVENTION: PROPHYLAXIS AND THERAPY  
NUMBER OF SEQUENCES: 184  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE  
CITY: NEW YORK  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/796,101  
FILING DATE: 05-FEB-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: KOMALSKI, THOMAS J.  
REGISTRATION NUMBER: 32,147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 764-5574  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3409 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-796-101-18

Query Match 2.3%; Score 38.2; DB 3; Length 3409;  
Best Local Similarity 45.7%; Pred. No. 0.25;  
Matches 133; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 579 ATCTCATTTTCTTCTGAGTCTCTGAGGCTGTGCTGCTGAGCCTGGAACTGCAGCA 638  
Db 758 ACCCTCTTCTCATCTGATCTCTCGCGCATGCGCCGCTAGTCATCACAATAGAGAGCA 817  
Qy 639 GCTGCCAGGTATGAGAGGCGCCAGAGTAATTAAGCCCTTGGGGCATTCGGTAC 698  
Db 818 GACTCTCAGAGATGCGCCCGCCAGAAATGACTGGCAAGACCTTCATGACATCTCTC 877

Qy 699 ATGTGTCGTACATTTAATGCGCTCATCTGACATGATGATTCATCTTGGCGCCAG 758  
Db 878 AATGCGGCGCTTCATTAACCTGATTAACCTGAGCTTGATTAACAGAGCCGCTGGCCAG 937  
Qy 759 CAATGACTATAGCTGGGCGAGGTGTTACTGTTATTTCAACAGAAATGAATGATCCT 818  
Db 938 CATGACACTAGTACCTTGTACTCATTAACATTTTTCACACATGATGAGATATC 997  
Qy 819 CCTGATCATCCATCTTTCGTCTCTCCATCTCTTCTTACATCA 869  
Db 998 CATAAATATGATCAATGCTGAGACACTTGTCTCTCTCATCCAA 1048

RESULT 13  
US-09-085-273-54  
Sequence 54, Application US/09085273  
Patent No. 6267965

GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
APPLICANT: Pincus, Steven E.  
APPLICANT: Cox, William I.  
APPLICANT: Kauffman, Elizabeth K.  
TITLE OF INVENTION: RECOMBINANT POXVIRUS - CYTOMEGALOVIRUS,  
TITLE OF INVENTION: COMPOSITIONS AND USES  
NUMBER OF SEQUENCES: 176  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/085,273  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/471,014  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer Esq., William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3409 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-085-273-54

Query Match 2.3%; Score 38.2; DB 3; Length 3409;  
Best Local Similarity 45.7%; Pred. No. 0.25;  
Matches 133; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 579 ATCTCATTTTCTTCTGAGTCTCTGAGGCTGTGCTGCTGAGCCTGGAACTGCAGCA 638  
Db 758 ACCCTCTTCTCATCTGATCTCTCGCGCATGCGCCGCTAGTCATCACAATAGAGAGCA 817  
Qy 639 GCTGCCAGGTATGAGAGGCGCCAGAGTAATTAAGCCCTTGGGGCATTCGGTAC 698  
Db 818 GACTCTCAGAGATGCGCCCGCCAGAAATGACTGGCAAGACCTTCATGACATCTCTC 877  
Qy 699 ATGTGTCGTACATTTAATGCGCTCATCTGACATGATGATTCATCTTGGCGCCAG 758

Db 878 AATGGGGGCTTCATTACACTGATTAACCTCAGGGCTGTGTTATCAGAGGCCGCTTGGCCAG 937  
Qy 759 CAAATGACTTATAGCTGGGGCAGTGTACTTCTTATTTCAACAGAAATTAATGATCCT 818  
Db 938 CATCAGACTAGTACCTTGTACTCAATTAACATTTGTTCCACATGTAGAGATATC 997  
Qy 819 CCGATCAATCCATCCTTTCCTCTCTCCATTTCTTCTTCTTCAACATCAA 869  
Db 998 CATAATATGTATGATCAATGTGGGTGAGACCTGTCTCTCTCCATCCAA 1048

RESULT 14  
US-09-916-963-54

Sequence 54, Application US/09916963  
Patent No. 6632438

## GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo  
Pancus, Steven E.  
Cox, William I.  
Kaufman, Elizabeth K.  
TITLE OF INVENTION: RECOMBINANT POXVIRUS - CYTOMEGALOVIRUS,  
COMPOSITIONS AND USES

NUMBER OF SEQUENCES: 176

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/916,963  
FILING DATE: 26-Jul-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,014  
FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Frommer Esq., William S.  
REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2720

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-0712  
TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:

LENGTH: 3409 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 54:  
US-09-916-963-54

Query Match 2.3%; Score 38.2; DB 4; Length 3409;  
Best Local Similarity 45.7%; Pred. No. 0.25;  
Matches 133; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

Qy 579 ATCCGATTTTCTTCTGGGCTCTGCGGTGGCTGTCGCTGAGCGTGGAACTGCAGAA 638  
Db 758 AGCTCTTCTCTATCTGACTCTCTCGGCGAATGGCCCGTAGTCAATCAACATGAGAGCA 817  
Qy 639 GCTGCCAGATTATGGAAGCGGCGCAAGTGAATATTAAGCCCTTTGGGGCATTCGGTAC 698  
Db 818 GACTCTCAGAGATCGGCCCGCAATGATGACTGGGCAAAACCTTCATGCAAGATCTCTC 877  
Qy 699 ATGTGTCGTACATTTAATGCGCCCTCATCTGGACTAGTAATGATCTCTTCCGTGCCAG 758

Db 878 AATGGGGGCTTCATTACACTGATTAACCTCAGGGCTGTGTTATCAGAGGCCGCTTGGCCAG 937  
Qy 759 CAAATGACTTATAGCTGGGGCAGTGTACTTCTTATTTCAACAGAAATTAATGATCCT 818  
Db 938 CATCAGACTAGTACCTTGTACTCAATTAACATTTGTTCCACATGTAGAGATATC 997  
Qy 819 CCGATCAATCCATCCTTTCCTCTCTCCATTTCTTCTTCTTCAACATCAA 869  
Db 998 CATAATATGTATGATCAATGTGGGTGAGACCTGTCTCTCTCCATCCAA 1048

RESULT 15  
US-10-204-708-49

Sequence 49, Application US/10204708  
Patent No. 6677731

## GENERAL INFORMATION:

APPLICANT: OLEK, Alexander  
APPLICANT: BRENNER, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
FILE REFERENCE: 5013.1012  
CURRENT APPLICATION NUMBER: US/10/204,708  
CURRENT FILING DATE: 2003-05-06

PRIOR APPLICATION NUMBER: PCT/EP01/03971  
PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: DE 10019056.8  
PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 98  
SEQ ID NO 49

LENGTH: 6306  
TYPE: DNA

ORGANISM: Artificial Sequence  
FEATURES:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-204-708-49

Query Match 2.3%; Score 38.2; DB 4; Length 6306;  
Best Local Similarity 48.8%; Pred. No. 0.4; Length 6306;  
Matches 103; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 1156 TTGAGAGCTTCATTAATTTTCTAGGAAGGTGTAGTGTGTTCACTGTTTGGAG 1215  
Db 2662 TTGGGAGAGTGAATTTGTAATAATTTATGTAATTGTAAGTTTATTTTAACTT 2721  
Qy 1216 GACTCATGCTTTTAACTAACAATCGGCGAATTCAGGTGGGAGAGCCCTGTTATGG 1275  
Db 2722 GAATGATAGGTTTAAAGATTATTTTATTTTAAATTAATGTAATTCGTTAGTTA 2781  
Qy 1276 TACGTTTCTTGGCTTCTAGTAGCCCATAGTTTATCTGTGTTGAACTGTGCTGG 1335  
Db 2782 TATTTTCTTGTATATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAAT 2841  
Qy 1336 ATGACATTTCTCTGTGTTTGTGCTGATCT 1366  
Db 2842 AAGTGTAAATTAATTTTGTGTTATTT 2872

Search completed: May 29, 2004, 07:46:53  
Job time: 150 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:22:44 / Search time 18 Seconds  
(without alignments)  
1287.290 Million cell updates/sec

Title: US-10-063-686-32

Perfect score: 2315  
Sequence: 1 MSGRDTLGLCTIALALSLA.....QDDKHSURNEGTLEQAIYR 445

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Query Match Length	ID	Description
1	126	5.4	372 1	CYB_TRYBO
2	119	5.1	523 1	NU4M_PROMI
3	118	5.1	497 1	NU4M_ACACA
4	112.5	4.9	348 1	GALR_MOUSE
5	110	4.8	495 1	NU4M_MARPO
6	109	4.7	501 1	TLCR_RICPR
7	108.5	4.7	346 1	GALR_RAT
8	108.5	4.7	608 1	VATT_BOBIBU
9	108	4.7	474 1	Y294_MYCCE
10	106	4.6	497 1	LMP2_EBV
11	106	4.6	677 1	ST21_ARATH
12	104.5	4.5	475 1	Y294_MYCPN
13	103	4.4	465 1	NP11_RAT
14	103	4.4	518 1	FL01_CRIGR
15	102.5	4.4	306 1	CTDA_BACSU
16	102	4.4	378 1	CKD6_MOUSE
17	102	4.4	422 1	RBN_CHLTE
18	102	4.4	433 1	NU4M_RHISA
19	102	4.4	446 1	DADR_HUMAN
20	102	4.4	446 1	DADR_YACMU
21	101.5	4.4	491 1	NU0M_RICPR
22	101.5	4.4	495 1	NU4M_WHAT
23	101	4.4	395 1	TRFR_CHICK
24	101	4.4	481 1	Y389_CLOPE
25	100.5	4.3	514 1	COX1_OSTNU
26	100.5	4.3	530 1	COX1_MERSE
27	100	4.3	386 1	NU4M_ARTSP
28	99.5	4.3	355 1	OPSG_CHICK
29	99.5	4.3	371 1	CML1_HUMAN
30	99	4.3	373 1	CML1_HUMAN
31	99	4.3	527 1	NU2M_ACACA
32	99	4.3	615 1	NU0L_BUCAP
33	98.5	4.3	394 1	TSGA_YERPE

34	98.5	4.3	562 1	ATKA_YERPE	08z96 yersinia pe
35	98.5	4.3	889 1	TRK2_YEAST	P28584 saccharomyc
36	98	4.2	318 1	MAAR_HUMAN	P33765 homo sapien
37	98	4.2	337 1	TR39_HUMAN	P35934 homo sapien
38	98	4.2	446 1	NU4M_DROME	P16931 drosophila
39	98	4.2	447 1	NU4M_APIIL	P34853 apis mellif
40	98	4.2	525 1	COX1_CAEL	P24893 caenorhabdi
41	98	4.2	663 1	CYOB_ECOLI	P18401 escherichia
42	97.5	4.2	250 1	ATP6_MYCTU	Q10591 mycobacteri
43	97.5	4.2	485 1	YIDL_ECOLI	P39276 escherichia
44	97.5	4.2	512 1	FL01_MOUSE	P41438 mus musculu
45	97.5	4.2	582 1	Y596_CAEL	Q09965 caenorhabdi

## ALIGNMENTS

RESULT 1  
CYB\_TRYBO STANDARD; PRI; 372 AA.  
AC Q33568; Q35988; 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Cytochrome b.  
GN MTCYB OR COB OR CYTB OR CYB.  
OS Trypanoplasma borreli.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Bodonidae; Trypanoplasma.  
OX NCBI\_Taxid=5710;  
RN [1]  
RP SEQUENCE FROM N.A., AND RNA EDITING.  
RC STRAIN=PG-0H;  
RX MEDLINE=95059046; PubMed=7969154;  
RA Maslov D.A., Simpson L.;  
RT "RNA editing and mitochondrial genomic organization in the cryptobid  
kinetoplastid protozoan Trypanoplasma borreli.",  
RL Mol. Cell. Biol. 14:8174-8182(1994).  
RN [2]  
RP SEQUENCE FROM N.A., AND RNA EDITING.  
RC STRAIN=TT-0H;  
RA Lukes J., Van den Burg J., Arts G.J., de Haan A., Oppendoes F.,  
Sloof P., Benne R.;  
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase  
complex (complex III or cytochrome b-c1 complex), which is a  
respiratory chain that generates an electrochemical potential  
coupled to ATP synthesis (By similarity).  
CC -!- COFACTOR: Binds two heme groups non-covalently. Heme 1 (or BL or  
b562) is low-potential and absorbs at about 562, and heme 2 (or BH  
or b566) is high-potential and absorbs at about 566 (By  
similarity).  
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,  
cytochrome c1 and the Rieske protein (By similarity).  
CC -!- RNA EDITING: Modified positions=Not applicable; Note=Some  
positions are modified by RNA editing via nucleotide insertion or  
deletion.  
CC -!- SIMILARITY: Belongs to the cytochrome b family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; U1482; AA65017.1; ALT\_SEQ.  
CC EMBL; U1484; AA73455.1; -.  
CC PIR; S52054; S52054.  
CC InterPro: IPR005798; Cytb b6 C.  
CC InterPro: IPR005797; Cytb b6 N.  
CC Pfam; PF00032; cytochrome\_b\_c1.1.

DR Pfam: PF00033; Cytochrome b N<sup>1</sup>; 1.  
 DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
 DR PROSITE; PS00193; CYTOCHROME\_B\_OO; 1.  
 KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;  
 KW Heme; RNA editing.  
 FT METAL 83 83 IRON 1 (HEME B562 AXIAL LIGAND).  
 FT METAL 97 97 IRON 2 (HEME B562 AXIAL LIGAND).  
 FT METAL 184 184 IRON 1 (HEME B562 AXIAL LIGAND).  
 FT METAL 198 198 IRON 2 (HEME B562 AXIAL LIGAND).  
 SQ SEQUENCE 372 AA; 44710 MW; DC4AF96804A0CDE6 CRC64;  
 Query Match 5.4%; Score 126; DB 1; Length 372;  
 Best Local Similarity 20.7%; Pred. No. 0.026;  
 Matches 92; Conservative 64; Mismatches 124; Indels 164; Gaps 24;  
 QY 10 LCIALALSLAMFTRFTTLLVHIFISLVILGILFVCGVLMWLYD-----YTND 61  
 DB 17 LCCLLISGDLFRYYGNGFNLGWMIALQI-----LVGIC--LSWFFRCIIPQNIWIFTL 68  
 QY 62 LSIETDERENMKCVGFAYSTGI--TAVLVLFVLAKRKLTVLFOITNKALISSAP 119  
 DB 69 IHFERP-----LGIIRSLHIFISLVFLYFLHITIKVIFCLIPDS----- 111  
 QY 120 FLIFQPLMTFAILIFFWLVMAVLLSLG-TAGAAQVMEGGQVEYK-----PLSGIRYMW 172  
 DB 112 ---MLVWFGFLIFILILAF--IGYTLPTCSMSYGLVFNILATILPLGI----- 160  
 QY 173 SYHLGLIWTSEI-----LACQMTIAG--VTCYFNKSKNDPDHFTL 216  
 DB 161 --YICQMWCSERINDFTLLKHSIHIFPVLFLIGAHFVLHYF-----L 206  
 QY 217 SLSLIL-----FFYHQTGVKXGSLISVARIPIRIIWMQNALKEQHGALSRVLF 267  
 DB 207 SSGDLDRPPFYERFFFLYYLRLDPLI--INLCFLIYI----- 247  
 QY 268 RCYCCEFWGLDKYHLHNONATTAINGTD-----FCTAKDAFKILSKNS 314  
 DB 248 ---CIYW---FFVHEHSMIIVDTLTKSDKILPEWFFLSFGFKLVDPKF----- 292  
 QY 315 SHFTSINCGRDIFILGKVLVVCFTVF---GGLMAFYNRAFOYVAVPLLVAFVATLV 370  
 DB 293 -----MGLFLF-----VLCFALFLFIANCILIFCYGSSLLM-MSLSLVLFY-VLC 337  
 QY 371 AHSFLSVETVIDALFLCPADVLE 394  
 DB 338 VGGFLSLY-----VVLCPFLMME 355  
 RESULT 2  
 NT4M PROWI STANDARD; PRT; 523 AA.  
 ID NT4M PROWI  
 AC Q37617;  
 DT 15-UTL-1998 (Rel. 36, Created)  
 DT 15-UTL-1998 (Rel. 36, Last sequence update)  
 DT 15-UTL-1998 (Rel. 36, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).  
 GN ND4 OR NAD4.  
 OS Prototheca wickerhamii.  
 OC Mitochondrion.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 OC Chlorellaceae; Prototheca.  
 OX NCBI\_TaxID=3111;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=263-11.  
 RX MEDLINE=94180393; PubMed=8133522;  
 RA Wolff G., Plante I., Lang B.F., Kueck U., Burger G.,  
 RT "Complete sequence of the mitochondrial DNA of the chlorophyte alga  
 RL J. Mol. Biol. 237:75-86(1994).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; U02970; AAD12635.1; -  
 DR PIR; T11916; T11916.  
 DR InterPro; IPR003918; NADHub oxd4.  
 DR InterPro; IPR001750; Oxidored\_q1.  
 DR Pfam; PF00361; oxidored\_q1; 1.  
 DR PRINTS; PRO1437; NOOXDRDTASE4.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 523 AA; 58945 MW; 2C558DD00935801D CRC64;  
 Query Match 5.1%; Score 119; DB 1; Length 523;  
 Best Local Similarity 18.1%; Pred. No. 0.12;  
 Matches 76; Conservative 63; Mismatches 163; Indels 118; Gaps 15;  
 QY 11 CIALALSLAMFTRFTTLLVHIFISLVILGILFVCGVLMWLYDNTDLSIEDTER 70  
 DB 151 CIAFLVLETMLTFVSVLDLILFYIFPESVLIEMFIIGV-WG-----SRER 196  
 QY 71 ENMKCVGFAYSTGITAVLVLFVLRKRLKTVLFOITNKALISSAPLLFOPLMTA 130  
 DB 197 KIRAAVOFLYTLFGSVLMLAILLYPQGTIDIMLVSPSEETROCILMLAFPASPA 256  
 QY 131 IL-----IFPW-----VLVAVLISLTAG--AAQVMEGGQVEYKLSGI 168  
 DB 257 VKYPMVPHVHMLBEAHEVAFTAGSVILAGILKLTGYGLRSIPLPFACIYFTPL 313  
 QY 169 RYMSYHLGLIWTSS-----EPLACQMTIAGAVVTCYFNKSKNDPDHFTLSL 219  
 DB 314 --IYMSVLAIVTSCITIRQIDLKILAVSSVAHNFITIGFSQNT----- 359  
 QY 220 SILFFYHQTGVKXGSLISVARIPIRIIWMQNALKEQHGALSRVLFCCYCCFWGLDK 279  
 DB 360 ---QG--IGSLLMI-----SHGLVSPALFLCVGVLY----- 387  
 QY 280 YLHLHNONATTAINGTDFCTSAKDAFKILSKNSHFTSINCF-----GDFIIFLG 331  
 DB 388 -----DRHKRLRLYYSGCGGTWPIFALL--FVFFMANISLPGTSFPGFELVFTG 437  
 QY 332 KVLVCEFTVFGGLMANNRAFOYVAVPLLVAFPIYLAH-SFLSVETVIDALFLCPA 390  
 DB 438 SYQNSFVAFCATGTGVLGAAYALMLCNRLI-----YGVSKDPDFTINTSDVVRREFFWA 492  
 RESULT 3  
 NT4M ACACA STANDARD; PRT; 497 AA.  
 ID NT4M ACACA  
 AC Q37375;  
 DT 15-UTL-1998 (Rel. 36, Created)  
 DT 15-UTL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).  
 GN ND4 OR NAD4.  
 OS Acanthamoeba castellanii (Amoebea).  
 OC Mitochondrion.  
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.  
 OX NCBI\_TaxID=5755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30010 / Neff.  
 RX MEDLINE=95147275; PubMed=7644823;  
 RA Burger G., Plante I., Loneragan K.M., Gray M.W.,  
 RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba  
 RL castellanii: complete sequence, gene content and genome  
 RT organization".  
 CC J. Mol. Biol. 245:522-537(1995).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: J12386; AAD11826.1; -  
DR PIR: S53834; S53834.  
DR InterPro: IPR003918; NADhub oxid4.  
DR InterPro: IPR001750; Oxidored\_g1.  
DR Pfam: PF00361; oxidored\_g1; 1.  
DR PRINTS: PR01437; NDOXRDTAS4.  
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
KW SEQUENCE 497 AA; 57768 MW; 9C0B637E65C0F3B5 CRC64;

Query Match 5.1%; Score 118; DB 1; Length 497;  
Best Local Similarity 19.5%; Pred. No. 0.13;  
Matches 90; Conservative 65; Mismatches 143; Indels 164; Gaps 22;

QY 11 CLIALALSLAMFTFFITTLVHIFISLVITG-----LLFVCGV 50  
DB 140 CVLTL-----VEFYIFESILIPMIFILGVGSSQKRIHAYQLFFYTLGLSLMLGI 193  
QY 51 LMMVYDYNDLSIEDTERENMKCVLGAIVSTGIVTLVLIIVLRKRIQLVTELQI 110  
DB 194 LVYSHIQTDRLVLTNIV-----SFYRQILLWASFF-- 226  
QY 111 TNKAISSAPFLLEQPLWTFALIFPW-----VLMVAVLLSLGTAGAQAQWEG 157  
DB 227 -----FAFCVKVPLEPFIH-----WLEPAHEAPTVGSVILAGLKLGTGL----- 269  
QY 156 GGVEXKPLSGIRYM-----NSYHLIGLWTFSEFLACQOMTIAQAVTCTYFNSSKNDP 211  
DB 270 -----LRFVPIPCDATYFPLPVYT-----LCLLGIITTCSTTRQVYLK 310  
QY 212 DHPILSLSLFPPYQGVTVKSPFLISVVRIPRIIWMVQNALKEQAGALSRYLFCY 271  
DB 311 KYIAVASY-----HMSFVILGLFTSNIGIGSVFNL-----SHGIVSSGLFRCIG 358  
QY 272 CCFWCLDK-----LHLNQAVYTTIANGDFCTSKADAPKILKSSSHFTSINCFDFTI 328  
DB 359 CVY-----DRYKRIKILVYSGLVSTMPI--FSLC-----FILLISNISPPGSSSFIGEPLI 408  
QY 329 FLG-----KVLVVCFTVFGGLMAFNRAFOVAVLPLLVAFAYLVASHFL-----S 376  
DB 409 LLGLPENHFAALITFSLI-----LTAVYSIMLVNRII--FNRLVYVYILRFSDFS 458  
QY 377 VFEYVLDLFLUCFAVDLETNDS-----SEKPYFMDQEFISF 413  
DB 459 KKEFVVGFI--CFITILFGLKSGYIISLEAPLYV---YLSF 496

RESUL 4  
ID GALT MOUSE STANDARD; PRT; 348 AA.  
AC P56479;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Galanin receptor type 1 (GAL1-R) (GALR1).  
GN GALR1 OR GALNR1 OR GALTNR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=97415411; PubMed=9271210;  
RA Wang S., He C., Maguire M.T., Clemmons A.L., Burrier R.E., Guzzi M.F.,  
RA Strader C.D., Parker E.M., Bayne M.L.;

RT "Genomic organization and functional characterization of the mouse  
RT GALR1 galanin receptor".  
RL FEBS Lett. 411:225-230(1997).  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=129;  
RX MEDLINE=98035873; PubMed=9367674;  
RA Jacoby A.S., Webb G.C., Liu M.L., Kofler B., Hort Y.J., Fathi Z.,  
RA Bottema C.D.K., Shine J., Iismaa T.P.;  
RT "Structural organization of the mouse and human GALR1 galanin  
RT receptor genes (galnr and GALTNR) and chromosomal localization of the  
RT mouse gene".  
RL Genomics 45:496-508(1997).

CC -1- FUNCTION: Receptor for the hormone galanin. The activity of this  
CC receptor is mediated by G proteins that inhibit adenylate cyclase  
CC activity.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- PM: THREE CYSTEINE RESIDUES ARE FOUND IN THE C-TERMINUS, AT LEAST  
CC ONE OF WHICH MAY BE PALMITOYLATED (BY SIMILARITY).  
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Y15004; CA75237.1; -  
DR EMBL: U90657; AAB87748.1; -  
DR EMBL: U90655; AAB87748.1; JOINED.  
DR EMBL: U90656; AAB87748.1; JOINED.  
DR MGD; MGI:1096364; Galr1.  
DR InterPro: IPR0020276; GPCR\_Rhodopsin.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCRHOPOPSN.  
DR PROSITE; PS00237; G\_PROTEIN\_RECP\_F1\_1; 1.  
DR PROSITE; PS00262; G\_PROTEIN\_RECP\_F1\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Multigene family; Lipoprotein; Palmitate.

FT DOMAIN 1 34  
FT TRANSSEM 35 55  
FT DOMAIN 56 70  
FT TRANSSEM 71 91  
FT DOMAIN 92 109  
FT TRANSSEM 110 131  
FT DOMAIN 132 151  
FT TRANSSEM 152 172  
FT DOMAIN 173 197  
FT TRANSSEM 198 218  
FT DOMAIN 219 247  
FT TRANSSEM 248 268  
FT DOMAIN 269 270  
FT TRANSSEM 271 291  
FT DOMAIN 292 348  
FT CARBOHYD 7 7  
FT CARBOHYD 12 12  
FT CARBOHYD 182 182  
FT DISULFID 106 186  
FT LIPID 319 319  
SQ SEQUENCE 348 AA; 39114 MW; 6F52D752BA19F9A CRC64;  
Query Match 4.9%; Score 112.5; DB 1; Length 348;  
Best Local Similarity 19.4%; Pred. No. 0.24;  
Matches 83; Conservative 43; Mismatches 129; Indels 173; Gaps 17;  
QY 33 VHFISLVILGLFPCGVLMWLYDYNDLSIEDTERENMKCVLGAIVSTGIVTLVAV 92  
DB 31 VENFITLVVFGILFAMGV-----GNSLVITV 57  
QY 93 LIFVLRKRIQLVTELQINRKAISSAPFLF-----QPLMTFAILL--FFWVLM 139

DB 58 IARSKPKRSTTNLF-ILNLSIADLAVLLFCIPQAVVAVLPTWVLAFCIKFIHYFFT 116  
QY 140 VAVLSLIGTAGAAY-----MEGGVEYKPLSGIRWMSYHILGLIMSEFLA 188  
DB 117 VSNLVSIFLLAMSDRYAIVHSKRSSSLRSKALLGVGFIWA----- 161  
QY 189 CQOMTITAGAVTCY-----FNRSKNDP-----PDHPILSSILFPGYQTVVKSFLIS 238  
DB 162 ---LSIIMASPVAYHQRLEPHRDSNOTFCWEQWPNK-----LHKKAYVCTFEVG 207  
QY 239 VVRIPRIIVMWNLMKEQHGALSRYLERCYCCFCWCLDKYLLHNLQNAVTTATNGTD 298  
DB 208 -----YLLPLLLITCF-CYAKVNLHLHK----- 229  
QY 299 FCTSAKDAFKILSKNSHFTS-INCFGDFIIFLAKVAVVCTFVGGMAFNRYAFOVA-- 356  
DB 230 -----LKNMSKKSSEASKKKT-----AQTVLVVVVVG--ISMLEPHVYHMAE 270  
QY 357 --AVPLLVAFPAVLVAH-----SEISV-FEYVLDALFLCFAVDLETNDGS 399  
DB 271 FGAFPLTPASFFFRITLHCLAVSNSSVNPDIYAFLSFNFRAYKQVFCVCHVDSBPSET 330  
QY 400 SEKPYEMD 407  
DB 331 KENSRMD 338

RESULT 5  
NU4M MARPO STANDARD; PRT; 495 AA.  
AC P26848;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).  
OS ND4 OR NAD4.  
GN Marchantia polymorpha (Liverwort).  
OC Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiaceae;  
OC Marchantiaceae; Marchantia.  
OX NCBI\_TaxId=3197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92114051; PubMed=1731062;  
RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,  
RA Akashi K., Kaneage T., Ogura Y., Kohchi T., Ohyama K.;  
RT "Gene organization deduced from the complete sequence of liverwort  
Marchantia polymorpha mitochondrial DNA. A primitive form of plant  
mitochondrial genome.";  
RL J. Mol. Biol. 223:1-7(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93247547; PubMed=8483448;  
RA Nozato N., Oda K., Yamato K., Ohta E., Takemura M., Akashi K.,  
RA Fukuzawa H., Ohyama K.;  
RT "Cotranscriptional expression of mitochondrial genes for subunits of  
NADH dehydrogenase, nad5, nad4, nad2, in Marchantia polymorpha.";  
RL Mol. Genet. 237:343-350(1993).  
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC EMBL, M68929; AAC09398.1; -  
CC PIR, S25942; S25942.  
CC InterPro, IPR003918; NADHnb oxrred4.

DR PFAM; PF00361; oxidored\_ql; 1.  
DR PRINTS; PR01437; NUXORDETAS54.  
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
SQ SEQUENCE 495 AA; 56311 MW; 0F75894D6C6AAED4 CRC64;  
Query Match 4.8%; Score 110; DB 1; Length 495;  
Best Local Similarity 18.9%; Pred. No. 0.52;  
Matches 78; Conservative 73; Mismatches 152; Indels 110; Gaps 16;  
QY 12 ILALSLAMFMFFRITTLVHIFSLVILG-----LFLVCGVL 51  
DB 130 LIAVFCSLDLILFYFESVLIIPMTIIGVGSRRKIRAAVQFLYTIMGSLFMILALIL 189  
QY 52 MWLYDYVNDLSIELTERENMKCVLGAIVSGITAVLVILFVLRKRIKLTVELFOIT 111  
DB 190 FIFQGTGTDLIQLTTE-----FSRRQILMIAFNASF 224  
QY 112 NKAISAPFLFQPLTFALILFWVLYAVVLSLTAG-----AAQVMEGGVEYKPLSG 167  
DB 225 SVKVPVPAVIMVPEAHVAPTAGSVTLAIGILIKGTGYFLRFSIPMPPEATLYFTP-- 281  
QY 168 IRYMMSYHLIGLIMTSEFLAACQMTITAGAVVTCYFNRSKNDPDPHPILSSILFVHQ 227  
DB 282 --FTYLSVIAIYTS--LTTIRQIDLKIIA-----YSSVA-----HM 316  
QY 228 GTVVKSPFLISVVRIPRIITVMQNALKEQHGALSRYLERCYCCFCWCLDRY---LML 284  
DB 317 NFVTIGMFSLNIGIGISILMLD-----SHGLVSSALFLCVGALY---DHKRIYVX 366  
QY 285 NQNAVTTTANGDPFCTSAKDAFKILSKNSHFTS-INCFGDFIIFLAKVAVVCTFV 340  
DB 367 YGGLVSTMPFSTIF-----LFTLANVSLPGTS-SFISGFLILVAGAFQNSLVATLAA 419  
QY 341 FGGLMAFNRYAFOVAVPLLVAFPA--YLVASHSLSVFETVLDALFLCFAV 391  
DB 420 LGWIL-----GAATSLVLYNRVVGNGKPFILKFSQNLNRREV---IFLPFIV 465

RESULT 6  
TLCC\_RICPR STANDARD; PRT; 501 AA.  
ID TLCC\_RICPR  
AC Q9ZD67;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE ADP/ATP carrier protein 3 (ADP/ATP translocase 3).  
GN TLCC OR TLG3 OR RP477.  
OS Rickettsia prowazekii.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
OX NCBI\_TaxId=782;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Madrid E;  
RX MEDLINE=99039499; PubMed=9823893;  
RA Anderson S.G.E., Zomrodipour A., Anderson J.O.,  
RA Scheritz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,  
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;  
RT "The genome sequence of Rickettsia prowazekii and the origin of  
mitochondria.";  
RL Nature 396:1133-1140(1998).  
CC -1- FUNCTION: PROVIDES THE RICKETTSIAL CELL WITH HOST ATP IN  
CC EXCHANGE FOR RICKETTSIAL ADP. THIS IS AN OBLIGATE EXCHANGE SYSTEM.  
CC THIS ENERGY ACQUIRING ACTIVITY IS AN IMPORTANT COMPONENT OF  
CC RICKETTSIAL PARASITISM.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: Belongs to the ADP/ATP translocase tlc family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way





```

FT CONFLICT 197 197 K -> N (IN REF. 2).
SQ SEQUENCE 474 AA; 52649 MW; 665273D90312D551 CRC64;
Query Match
Best Local Similarity 4.7%; Score 108; DB 1; Length 474;
Matches 88; Conservative 55; Mismatches 132; Indels 140; Gaps 20;
QY 37 ISVILGL--FVCGVLMW-----LYDYNDLSIELDTERENKCVIGFATVSTG 85
DB 22 LALIIIGAIIDVFYIAAPYVKNVVPNIHLITGITED-----EVAIVTSIIQVTLAQ 74
QY 86 ITAVLVLIIVLKRRIKLTVELFQITNKAIS-----SAPFLIFQPLMT- 129
DB 75 LPGGFLTRSSSKLFLSA-----ITTGALTFWLANILTKNOQSHDALFIOYCVIWM 130
QY 130 ---AIIFFVVLWVAV-----LISLTAGAAQVMEGQVEYPLSGIRYMSYHLI 177
DB 131 GISTSLIFWPLWKLASQATKENQALGFGIGQAA-----NGIMGLIFIFI 177
QY 178 GLTWSEFILAQQMTIAGAVTCYFRSKNDPPHP-----ILSSSLIFPY 225
DB 178 ALIITISF-----YPSGSENDSKFPAAVFIAMLVITGFTVLFV 221
QY 226 HQTIVVKS--FLISVVR--IPRIIWMQNALKEQGHALSRYLFCRCYCCFWCLDKYLL 282
DB 222 KEKPIEKQSGTTLVSPFRNLNQLVTLKMKMLLSFFLMGMVYVQSTFA-----YILL 275
QY 283 HLNQNVY-----TTTLING--TDFTSADAF--KILSKNSHFTSINCQDPIIF- 329
DB 276 QMLQNAFLAPVVLVTVIGIRITVARSAYSVCRLADCKS-----YIIFLMTC 325
QY 330 --LGKLVVCFVTVFGGLMAFNVRARQWAVPLLVAFAYLVASHFSVEFTVL 382
DB 326 TVLGIVFVLAFLILGIVQINSAN-----ITLITF--SSILYIFIGIL 365

RESULT 10
LMP2_EBV
ID LMP2_EBV STANDARD; PRT; 497 AA.
AC P13285;
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Gene terminal protein (Membrane protein LMP-2A/LMP-2B).
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8283646; PubMed=2840285;
RA Lau G., Perricaudet M., Farrell P.J.;
RT "A spliced Epstein-Barr virus gene expressed in immortalized
RT lymphocytes is created by circularization of the linear viral
RT genome.";
RL EMBO J. 7:769-774 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=8905024; PubMed=2536113;
RA Sample J., Liebowitz D., Kieff E.;
RT "Two related Epstein-Barr virus membrane proteins are encoded by
RT separate genes.";
RL J Virol. 63:933-937 (1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=8427067; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Watfull B.G., Watson G.S., Weller J.C., Wilson
RA "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211 (1984).
CC -1- FUNCTION: MAY BE IMPORTANT FOR IMMORTALIZATION OF CELLS BY EBV.
CC -1- SUBCELLULAR LOCATION: Membrane (Probable).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M24212; AAA45887.1; -.
DR EMBL; Y00835; CAA68762.1; -.
DR EMBL; V01555; -. NOT ANNOTATED_CDS.
DR PIR; A30178; WMBELM.
KW Transmembrane.
FT CHAIN 1 497 MEMBRANE PROTEIN LMP-2A.
FT TRANSMEM 120 497 MEMBRANE PROTEIN LMP-2B.
FT TRANSMEM 122 141 POTENTIAL.
FT TRANSMEM 150 168 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 208 235 POTENTIAL.
FT TRANSMEM 242 259 POTENTIAL.
FT TRANSMEM 267 288 POTENTIAL.
FT TRANSMEM 300 316 POTENTIAL.
FT TRANSMEM 321 339 POTENTIAL.
FT TRANSMEM 355 373 POTENTIAL.
FT TRANSMEM 392 411 POTENTIAL.
FT TRANSMEM 419 443 POTENTIAL.
FT TRANSMEM 450 470 POTENTIAL.
SQ SEQUENCE 497 AA; 53011 MW; F4DC9B83C1FD83F1 CRC64;
Query Match
Best Local Similarity 4.6%; Score 106; DB 1; Length 497;
Matches 72; Conservative 49; Mismatches 126; Indels 110; Gaps 13;
QY 6 TILGCIATLALSLAMMFTFRFTT---LIVHIFISLVIIIGLLFVCGVLMWLYYDITNDL 62
DB 150 TVTATATGLALSLLILAAVASSYAAQKLTPTVTVLTAVTPEAIC-LTWRIIDPPNSL 208
QY 63 SIEDTERENKCVLGFATVSTGITAVLVLIFFVLRKRRLTV-----ELFOITNK 113
DB 209 LFLA-----LAAAGIGQIYVLMVLILAVRRMR-RLTVCGGIMFLACVLIVDA 261
QY 114 AISAAPLTFQPLMTFAFLILFFFWLVAVAVLSIGTGAQVMEGQVEY--KLSGIRN 171
DB 262 VLQSPILGAVTVVSMTLILAFVLMWSFGGIGTIGALTLTAAALIASLITGLTMD 321
QY 172 WSYHLIGLWTSFIIACQMT-----IAGAVVTCYFN 205
DB 322 TTMFLMLMTLVVLVILCSSCSPLSKILLARLFYALALLILASALTAGSILDTNKR 381
QY 206 SKNDPPHP-----ILSSSLIFPYQG-----TVKGSFLIS 238
DB 382 SLSTSEFIPULFQMLLIIVAGILFILATLTWSSGNRTYGPVFMCLGGLTWAGAVVWL 441
QY 239 VVRIPRIIWMQNALKEQGHALSRYL-----FRCC-YCCFWCL 277
DB 442 V-----MNTL-----LSAMILTAGLFLILGIFALGEGVIRCCRYCYCL 481
RESULT 11
ST21_ARATH
ID ST21_ARATH STANDARD; PRT; 677 AA.
AC O04722;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfate transporter 2.1 (AST68).
GN SUTR2.1 OR ATSG10180 OR T3IP16_170.
OS Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;

```

[1]  
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND INDUCTION.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=98021500; PubMed=9380766;  
 RA Takahashi H., Yamazaki M., Sasakura N., Watanabe A., Leustek T.,  
 RA de Almeida Engler J., Engler G., Van Montagu M., Saito K.;  
 RT "Regulation of sulfur assimilation in higher plants: a sulfate  
 transporter induced in sulfate-starved roots plays a central role in  
 Arabidopsis thaliana.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11102-11107(1997).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016721; PubMed=11130714;  
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,  
 RA Kohara M., Matsunoto M., Matsuno A., Muraki A., Nakayama S.,  
 RA Nakazaki N., Naito K., Okumura S., Shino S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,  
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,  
 RA Stoeckling T., Pepin K., Spieth J., Sekhon J., Armstrong J., Becker M.,  
 RA Belter B., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,  
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar B., Latreille P.,  
 RA Leonard S., Meyer R., Mulvaney E., Ozeresky P., Riley A., Strommatt C.,  
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedina N.,  
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vill D., Baker J.,  
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,  
 RA Martensen R., McComble W.R., Wilson R.K., Murphy G., Bancroft I.,  
 RA Volkart G., Wandut R., Duesetnoet A., Stekema W., Pohl T.,  
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,  
 RA Langham S.-A., McCullagh B., Robben J., Grymoprez B., Zimmermann W.,  
 RA Ranspenger U., Medler H., Balke K., Medler E., Peters S.,  
 RA van Steveren M., Dirke W., Moolman P., Klein lankhorst R.,  
 RA Weitzenecker T., Bothe G., Rose M., Hauf J., Bernier S., Hempel S.,  
 RA Bents O., Lemcke K., Kolesov G., Villarroel R., Giesen U., Ardiles W.,  
 RA Scheller C., Zaccaria P., Mewes H.-W., Beyan K.F.X., Rüd S., Schoof H.,  
 RA "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 408:823-826(2000).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Phan P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Tortum M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Anstett Y.,  
 RA Arakawa T., Bann J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Eju N., Goldsmith A.D., Gujral M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,  
 RA Khan S., Kosemura E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamliya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 genome.";  
 RL Science 302:842-846(2003).  
 RP FUNCTION, TISSUE SPECIFICITY, AND INDUCTION.  
 RX MEDLINE=20387013; PubMed=10929111;  
 RA Takahashi H., Melanabe-Takahashi A., Smith F.W., Blake-Kalif M.,  
 RA Hawkesford M.J., Saito K.;  
 RT "The roles of three functional sulphate transporters involved in  
 uptake and translocation of sulphate in Arabidopsis thaliana.";  
 RL Plant J. 23:171-182(2000).  
 RP FUNCTION: Low-affinity H(+)/sulfate cotransporter that may be  
 involved in root-to-shoot translocation of sulfate. Plays a  
 central role in the regulation of sulfate assimilation.  
 CC INVOLVED IN ROOT-TO-SHOOT TRANSLLOCATION OF SULFATE ASSIMILATION.  
 CC SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC TISSUE SPECIFICITY: Expressed in root cap, central cylinder of

CC roots and in vascular tissues of leaves.  
 CC -1- INDUCTION: In roots by sulfate starvation or after selenate  
 CC treatment.  
 CC -1- SIMILARITY: Belongs to the SLC26A/Sulph transporter (TC 2.A.53)  
 CC family.  
 CC -1- SIMILARITY: Contains 1 STAS domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; AB003590; BAA20084.1; -  
 CC EMBL; AB003591; BAA20085.1; -  
 CC EMBL; AL356332; CAB92059.1; -  
 CC EMBL; AY062846; AL332624.1; -  
 CC EMBL; AY093335; AAM13334.1; -  
 CC FTR; T50022; T50022.  
 CC InterPro; IPR002645; STAS.  
 CC InterPro; IPR001902; Sulph\_transp.  
 CC Pfam; PF01740; STAS; 1.  
 CC Pfam; PF00916; Sulfate\_transp; 1.  
 CC TIGRfam; TIGR00815; sulp; 1.  
 CC PROSITE; PS01130; SLC26A; 1.  
 CC PROSITE; PS0801; STAS; 1.  
 CC TRANSPORT; Symport; Sulfate transport; Transmembrane; Glycoprotein;  
 CC Multigene family.  
 CC DOMAIN 1 118  
 CC TRANSMEM 119 139  
 CC DOMAIN 140 141  
 CC TRANSMEM 142 142  
 CC TRANSMEM 143 143  
 CC TRANSMEM 144 144  
 CC TRANSMEM 145 145  
 CC TRANSMEM 146 146  
 CC TRANSMEM 147 147  
 CC TRANSMEM 148 148  
 CC TRANSMEM 149 149  
 CC TRANSMEM 150 150  
 CC TRANSMEM 151 151  
 CC TRANSMEM 152 152  
 CC TRANSMEM 153 153  
 CC TRANSMEM 154 154  
 CC TRANSMEM 155 155  
 CC TRANSMEM 156 156  
 CC TRANSMEM 157 157  
 CC TRANSMEM 158 158  
 CC TRANSMEM 159 159  
 CC TRANSMEM 160 160  
 CC TRANSMEM 161 161  
 CC TRANSMEM 162 162  
 CC TRANSMEM 163 163  
 CC TRANSMEM 164 164  
 CC TRANSMEM 165 165  
 CC TRANSMEM 166 166  
 CC TRANSMEM 167 167  
 CC TRANSMEM 168 168  
 CC TRANSMEM 169 169  
 CC TRANSMEM 170 170  
 CC TRANSMEM 171 171  
 CC TRANSMEM 172 172  
 CC TRANSMEM 173 173  
 CC TRANSMEM 174 174  
 CC TRANSMEM 175 175  
 CC TRANSMEM 176 176  
 CC TRANSMEM 177 177  
 CC TRANSMEM 178 178  
 CC TRANSMEM 179 179  
 CC TRANSMEM 180 180  
 CC TRANSMEM 181 181  
 CC TRANSMEM 182 182  
 CC TRANSMEM 183 183  
 CC TRANSMEM 184 184  
 CC TRANSMEM 185 185  
 CC TRANSMEM 186 186  
 CC TRANSMEM 187 187  
 CC TRANSMEM 188 188  
 CC TRANSMEM 189 189  
 CC TRANSMEM 190 190  
 CC TRANSMEM 191 191  
 CC TRANSMEM 192 192  
 CC TRANSMEM 193 193  
 CC TRANSMEM 194 194  
 CC TRANSMEM 195 195  
 CC TRANSMEM 196 196  
 CC TRANSMEM 197 197  
 CC TRANSMEM 198 198  
 CC TRANSMEM 199 199  
 CC TRANSMEM 200 200  
 CC TRANSMEM 201 201  
 CC TRANSMEM 202 202  
 CC TRANSMEM 203 203  
 CC TRANSMEM 204 204  
 CC TRANSMEM 205 205  
 CC TRANSMEM 206 206  
 CC TRANSMEM 207 207  
 CC TRANSMEM 208 208  
 CC TRANSMEM 209 209  
 CC TRANSMEM 210 210  
 CC TRANSMEM 211 211  
 CC TRANSMEM 212 212  
 CC TRANSMEM 213 213  
 CC TRANSMEM 214 214  
 CC TRANSMEM 215 215  
 CC TRANSMEM 216 216  
 CC TRANSMEM 217 217  
 CC TRANSMEM 218 218  
 CC TRANSMEM 219 219  
 CC TRANSMEM 220 220  
 CC TRANSMEM 221 221  
 CC TRANSMEM 222 222  
 CC TRANSMEM 223 223  
 CC TRANSMEM 224 224  
 CC TRANSMEM 225 225  
 CC TRANSMEM 226 226  
 CC TRANSMEM 227 227  
 CC TRANSMEM 228 228  
 CC TRANSMEM 229 229  
 CC TRANSMEM 230 230  
 CC TRANSMEM 231 231  
 CC TRANSMEM 232 232  
 CC TRANSMEM 233 233  
 CC TRANSMEM 234 234  
 CC TRANSMEM 235 235  
 CC TRANSMEM 236 236  
 CC TRANSMEM 237 237  
 CC TRANSMEM 238 238  
 CC TRANSMEM 239 239  
 CC TRANSMEM 240 240  
 CC TRANSMEM 241 241  
 CC TRANSMEM 242 242  
 CC TRANSMEM 243 243  
 CC TRANSMEM 244 244  
 CC TRANSMEM 245 245  
 CC TRANSMEM 246 246  
 CC TRANSMEM 247 247  
 CC TRANSMEM 248 248  
 CC TRANSMEM 249 249  
 CC TRANSMEM 250 250  
 CC TRANSMEM 251 251  
 CC TRANSMEM 252 252  
 CC TRANSMEM 253 253  
 CC TRANSMEM 254 254  
 CC TRANSMEM 255 255  
 CC TRANSMEM 256 256  
 CC TRANSMEM 257 257  
 CC TRANSMEM 258 258  
 CC TRANSMEM 259 259  
 CC TRANSMEM 260 260  
 CC TRANSMEM 261 261  
 CC TRANSMEM 262 262  
 CC TRANSMEM 263 263  
 CC TRANSMEM 264 264  
 CC TRANSMEM 265 265  
 CC TRANSMEM 266 266  
 CC TRANSMEM 267 267  
 CC TRANSMEM 268 268  
 CC TRANSMEM 269 269  
 CC TRANSMEM 270 270  
 CC TRANSMEM 271 271  
 CC TRANSMEM 272 272  
 CC TRANSMEM 273 273  
 CC TRANSMEM 274 274  
 CC TRANSMEM 275 275  
 CC TRANSMEM 276 276  
 CC TRANSMEM 277 277  
 CC TRANSMEM 278 278  
 CC TRANSMEM 279 279  
 CC TRANSMEM 280 280  
 CC TRANSMEM 281 281  
 CC TRANSMEM 282 282  
 CC TRANSMEM 283 283  
 CC TRANSMEM 284 284  
 CC TRANSMEM 285 285  
 CC TRANSMEM 286 286  
 CC TRANSMEM 287 287  
 CC TRANSMEM 288 288  
 CC TRANSMEM 289 289  
 CC TRANSMEM 290 290  
 CC TRANSMEM 291 291  
 CC TRANSMEM 292 292  
 CC TRANSMEM 293 293  
 CC TRANSMEM 294 294  
 CC TRANSMEM 295 295  
 CC TRANSMEM 296 296  
 CC TRANSMEM 297 297  
 CC TRANSMEM 298 298  
 CC TRANSMEM 299 299  
 CC TRANSMEM 300 300  
 CC TRANSMEM 301 301  
 CC TRANSMEM 302 302  
 CC TRANSMEM 303 303  
 CC TRANSMEM 304 304  
 CC TRANSMEM 305 305  
 CC TRANSMEM 306 306  
 CC TRANSMEM 307 307  
 CC TRANSMEM 308 308  
 CC TRANSMEM 309 309  
 CC TRANSMEM 310 310  
 CC TRANSMEM 311 311  
 CC TRANSMEM 312 312  
 CC TRANSMEM 313 313  
 CC TRANSMEM 314 314  
 CC TRANSMEM 315 315  
 CC TRANSMEM 316 316  
 CC TRANSMEM 317 317  
 CC TRANSMEM 318 318  
 CC TRANSMEM 319 319  
 CC TRANSMEM 320 320  
 CC TRANSMEM 321 321  
 CC TRANSMEM 322 322  
 CC TRANSMEM 323 323  
 CC TRANSMEM 324 324  
 CC TRANSMEM 325 325  
 CC TRANSMEM 326 326  
 CC TRANSMEM 327 327  
 CC TRANSMEM 328 328  
 CC TRANSMEM 329 329  
 CC TRANSMEM 330 330  
 CC TRANSMEM 331 331  
 CC TRANSMEM 332 332  
 CC TRANSMEM 333 333  
 CC TRANSMEM 334 334  
 CC TRANSMEM 335 335  
 CC TRANSMEM 336 336  
 CC TRANSMEM 337 337  
 CC TRANSMEM 338 338  
 CC TRANSMEM 339 339  
 CC TRANSMEM 340 340  
 CC TRANSMEM 341 341  
 CC TRANSMEM 342 342  
 CC TRANSMEM 343 343  
 CC TRANSMEM 344 344  
 CC TRANSMEM 345 345  
 CC TRANSMEM 346 346  
 CC TRANSMEM 347 347  
 CC TRANSMEM 348 348  
 CC TRANSMEM 349 349  
 CC TRANSMEM 350 350  
 CC TRANSMEM 351 351  
 CC TRANSMEM 352 352  
 CC TRANSMEM 353 353  
 CC TRANSMEM 354 354  
 CC TRANSMEM 355 355  
 CC TRANSMEM 356 356  
 CC TRANSMEM 357 357  
 CC TRANSMEM 358 358  
 CC TRANSMEM 359 359  
 CC TRANSMEM 360 360  
 CC TRANSMEM 361 361  
 CC TRANSMEM 362 362  
 CC TRANSMEM 363 363  
 CC TRANSMEM 364 364  
 CC TRANSMEM 365 365  
 CC TRANSMEM 366 366  
 CC TRANSMEM 367 367  
 CC TRANSMEM 368 368  
 CC TRANSMEM 369 369  
 CC TRANSMEM 370 370  
 CC TRANSMEM 371 371  
 CC TRANSMEM 372 372  
 CC TRANSMEM 373 373  
 CC TRANSMEM 374 374  
 CC TRANSMEM 375 375  
 CC TRANSMEM 376 376  
 CC TRANSMEM 377 377  
 CC TRANSMEM 378 378  
 CC TRANSMEM 379 379  
 CC TRANSMEM 380 380  
 CC TRANSMEM 381 381  
 CC TRANSMEM 382 382  
 CC TRANSMEM 383 383  
 CC TRANSMEM 384 384  
 CC TRANSMEM 385 385  
 CC TRANSMEM 386 386  
 CC TRANSMEM 387 387  
 CC TRANSMEM 388 388  
 CC TRANSMEM 389 389  
 CC TRANSMEM 390 390  
 CC TRANSMEM 391 391  
 CC TRANSMEM 392 392  
 CC TRANSMEM 393 393  
 CC TRANSMEM 394 394  
 CC TRANSMEM 395 395  
 CC TRANSMEM 396 396  
 CC TRANSMEM 397 397  
 CC TRANSMEM 398 398  
 CC TRANSMEM 399 399  
 CC TRANSMEM 400 400  
 CC TRANSMEM 401 401  
 CC TRANSMEM 402 402  
 CC TRANSMEM 403 403  
 CC TRANSMEM 404 404  
 CC TRANSMEM 405 405  
 CC TRANSMEM 406 406  
 CC TRANSMEM 407 407  
 CC TRANSMEM 408 408  
 CC TRANSMEM 409 409  
 CC TRANSMEM 410 410  
 CC TRANSMEM 411 411  
 CC TRANSMEM 412 412  
 CC TRANSMEM 413 413  
 CC TRANSMEM 414 414  
 CC TRANSMEM 415 415  
 CC TRANSMEM 416 416  
 CC TRANSMEM 417 417  
 CC TRANSMEM 418 418  
 CC TRANSMEM 419 419  
 CC TRANSMEM 420 420  
 CC TRANSMEM 421 421  
 CC TRANSMEM 422 422  
 CC TRANSMEM 423 423  
 CC TRANSMEM 424 424  
 CC TRANSMEM 425 425  
 CC TRANSMEM 426 426  
 CC TRANSMEM 427 427  
 CC TRANSMEM 428 428  
 CC TRANSMEM 429 429  
 CC TRANSMEM 430 430  
 CC TRANSMEM 431 431  
 CC TRANSMEM 432 432  
 CC TRANSMEM 433 433  
 CC TRANSMEM 434 434  
 CC TRANSMEM 435 435  
 CC TRANSMEM 436 436  
 CC TRANSMEM 437 437  
 CC TRANSMEM 438 438  
 CC TRANSMEM 439 439  
 CC TRANSMEM 440 440  
 CC TRANSMEM 441 441  
 CC TRANSMEM 442 442  
 CC TRANSMEM 443 443  
 CC TRANSMEM 444 444  
 CC TRANSMEM 445 445  
 CC TRANSMEM 446 446  
 CC TRANSMEM 447 447  
 CC TRANSMEM 448 448  
 CC TRANSMEM 449 449  
 CC TRANSMEM 450 450  
 CC TRANSMEM 451 451  
 CC TRANSMEM 452 452  
 CC TRANSMEM 453 453  
 CC TRANSMEM 454 454  
 CC TRANSMEM 455 455  
 CC TRANSMEM 456 456  
 CC TRANSMEM 457 457  
 CC TRANSMEM 458 458  
 CC TRANSMEM 459 459  
 CC TRANSMEM 460 460  
 CC TRANSMEM 461 461  
 CC TRANSMEM 462 462  
 CC TRANSMEM 463 463  
 CC TRANSMEM 464 464  
 CC TRANSMEM 465 465  
 CC TRANSMEM 466 466  
 CC TRANSMEM 467 467  
 CC TRANSMEM 468 468  
 CC TRANSMEM 469 469  
 CC TRANSMEM 470 470  
 CC TRANSMEM 471 471  
 CC TRANSMEM 472 472  
 CC TRANSMEM 473 473  
 CC TRANSMEM 474 474  
 CC TRANSMEM 475 475  
 CC TRANSMEM 476 476  
 CC TRANSMEM 477 477  
 CC TRANSMEM 478 478  
 CC TRANSMEM 479 479  
 CC TRANSMEM 480 480  
 CC TRANSMEM 481 481  
 CC TRANSMEM 482 482  
 CC TRANSMEM 483 483  
 CC TRANSMEM 484 484  
 CC TRANSMEM 485 485  
 CC TRANSMEM 486 486  
 CC TRANSMEM 487 487  
 CC TRANSMEM 488 488  
 CC TRANSMEM 489 489  
 CC TRANSMEM 490 490  
 CC TRANSMEM 491 491  
 CC TRANSMEM 492 492  
 CC TRANSMEM 493 493  
 CC TRANSMEM 494 494  
 CC TRANSMEM 495 495  
 CC TRANSMEM 496 496  
 CC TRANSMEM 497 497  
 CC TRANSMEM 498 498  
 CC TRANSMEM 499 499  
 CC TRANSMEM 500 500  
 CC TRANSMEM 501 501  
 CC TRANSMEM 502 502  
 CC TRANSMEM 503 503  
 CC TRANSMEM 504 504  
 CC TRANSMEM 505 505  
 CC TRANSMEM 506 506  
 CC TRANSMEM 507 507  
 CC TRANSMEM 508 508  
 CC TRANSMEM 509 509  
 CC TRANSMEM 510 510  
 CC TRANSMEM 511 511  
 CC TRANSMEM 512 512  
 CC TRANSMEM 513 513  
 CC TRANSMEM 514 514  
 CC TRANSMEM 515 515  
 CC TRANSMEM 516 516  
 CC TRANSMEM 517 517  
 CC TRANSMEM 518 518  
 CC TRANSMEM 519 519  
 CC TRANSMEM 520 520  
 CC TRANSMEM 521 521  
 CC TRANSMEM 522 522  
 CC TRANSMEM 523 523  
 CC TRANSMEM 524 524  
 CC TRANSMEM 525 525  
 CC TRANSMEM 526 526  
 CC TRANSMEM 527 527  
 CC TRANSMEM 528 528  
 CC TRANSMEM 529 529  
 CC TRANSMEM 530 530  
 CC TRANSMEM 531 531  
 CC TRANSMEM 532 532  
 CC TRANSMEM 533 533  
 CC TRANSMEM 534 534  
 CC TRANSMEM 535 535  
 CC TRANSMEM 536 536  
 CC TRANSMEM 537 537  
 CC TRANSMEM 538 538  
 CC TRANSMEM 539 539  
 CC TRANSMEM 540 540  
 CC TRANSMEM 541 541  
 CC TRANSMEM 542 542  
 CC TRANSMEM 543 543  
 CC TRANSMEM 544 544  
 CC TRANSMEM 545 545  
 CC TRANSMEM 546 546  
 CC TRANSMEM 547 547  
 CC TRANSMEM 548 548  
 CC TRANSMEM 549 549  
 CC TRANSMEM 550 550  
 CC TRANSMEM 551 551  
 CC TRANSMEM 552 552  
 CC TRANSMEM 553 553  
 CC TRANSMEM 554 554  
 CC TRANSMEM 555 555  
 CC TRANSMEM 556 556  
 CC TRANSMEM 557 557  
 CC TRANSMEM 558 558  
 CC TRANSMEM 559 559  
 CC TRANSMEM 560 560  
 CC TRANSMEM 561 561  
 CC TRANSMEM 562 562  
 CC TRANSMEM 563 563  
 CC TRANSMEM 564 564  
 CC TRANSMEM 565 565  
 CC TRANSMEM 566 566  
 CC TRANSMEM 567 567  
 CC TRANSMEM 568 568  
 CC TRANSMEM 569 569  
 CC TRANSMEM 570 570  
 CC TRANSMEM 571 571  
 CC TRANSMEM 572 572  
 CC TRANSMEM 573 573  
 CC TRANSMEM 574 574  
 CC TRANSMEM 575 575  
 CC TRANSMEM 576 576  
 CC TRANSMEM 577 577  
 CC TRANSMEM 578 578  
 CC TRANSMEM 579 579  
 CC TRANSMEM 580 580  
 CC TRANSMEM 581 581  
 CC TRANSMEM 582 582  
 CC TRANSMEM 583 583  
 CC TRANSMEM 584 584  
 CC TRANSMEM 585 585  
 CC TRANSMEM 586 586  
 CC TRANSMEM 587 587  
 CC TRANSMEM 588 588  
 CC TRANSMEM 589 589  
 CC TRANSMEM 590 590  
 CC TRANSMEM 591 591  
 CC TRANSMEM 592 592  
 CC TRANSMEM 593 593  
 CC TRANSMEM 594 594  
 CC TRANSMEM 595 595  
 CC TRANSMEM 596 596  
 CC TRANSMEM 597 597  
 CC TRANSMEM 598 598  
 CC TRANSMEM 599 599  
 CC TRANSMEM 600 600  
 CC TRANSMEM 601 601  
 CC TRANSMEM 602 602  
 CC TRANSMEM 603 603  
 CC TRANSMEM 604 604  
 CC TRANSMEM 605 605  
 CC TRANSMEM 606 606  
 CC TRANSMEM 607 607  
 CC TRANSMEM 608 608  
 CC TRANSMEM 609 609  
 CC TRANSMEM 610 610  
 CC TRANSMEM 611 611  
 CC TRANSMEM 612 612  
 CC TRANSMEM 613 613  
 CC TRANSMEM 614 614  
 CC TRANSMEM 615 615  
 CC TRANSMEM 616 616  
 CC TRANSMEM 617 617  
 CC TRANSMEM 618 618  
 CC TRANSMEM 619 619  
 CC TRANSMEM 620 620  
 CC TRANSMEM 621 621  
 CC TRANSMEM 622 622  
 CC TRANSMEM 623 623  
 CC TRANSMEM 624 624  
 CC TRANSMEM 625 625  
 CC TRANSMEM 626 626  
 CC TRANSMEM 627 627  
 CC TRANSMEM 628 628  
 CC TRANSMEM 629 629  
 CC TRANSMEM 630 630  
 CC TRANSMEM 631 631  
 CC TRANSMEM 632 632  
 CC TRANSMEM 633 633  
 CC TRANSMEM 634 634  
 CC TRANSMEM 635 635  
 CC TRANSMEM 636 636  
 CC TRANSMEM 637 637  
 CC TRANSMEM 638 638  
 CC TRANSMEM 639 639  
 CC TRANSMEM 640 640  
 CC TRANSMEM 641 641  
 CC TRANSMEM 642 642  
 CC TRANSMEM 643 643  
 CC TRANSMEM 644 644  
 CC TRANSMEM 645 645  
 CC TRANSMEM 646 646  
 CC TRANSMEM 647 647  
 CC TRANSMEM 648 648  
 CC TRANSMEM 649 649  
 CC TRANSMEM 650 650  
 CC TRANSMEM 651 651  
 CC TRANSMEM 652 652  
 CC TRANSMEM 653 653  
 CC TRANSMEM 654 654  
 CC TRANSMEM 655 655  
 CC TRANSMEM 656 656  
 CC TRANSMEM 657 657  
 CC TRANSMEM 658 658  
 CC TRANSMEM 659 659  
 CC TRANSMEM 660 660  
 CC TRANSMEM 661 661  
 CC TRANSMEM 662 662  
 CC TRANSMEM 663 663  
 CC TRANSMEM 664 664  
 CC TRANSMEM 665 665  
 CC TRANSMEM 666 666  
 CC TRANSMEM 667 667  
 CC TRANSMEM 668 668  
 CC TRANSMEM 669 669  
 CC TRANSMEM 670 670  
 CC TRANSMEM 671 671  
 CC TRANSMEM 672 672  
 CC TRANSMEM 673 673  
 CC TRANSMEM 674 674  
 CC TRANSMEM 675 675  
 CC TRANSMEM 676 676  
 CC TRANSMEM 677 677  
 CC TRANSMEM 678 678  
 CC TRANSMEM 679 679  
 CC TRANSMEM 680 680  
 CC TRANSMEM 681 681  
 CC TRANSMEM 682 682  
 CC TRANSMEM 683 683  
 CC TRANSMEM 684 684  
 CC TRANSMEM 685 685  
 CC TRANSMEM 686 686  
 CC TRANSMEM 687 687  
 CC TRANSMEM 688 688  
 CC TRANSMEM 689 689  
 CC TRANSMEM 690 690  
 CC TRANSMEM 691 691  
 CC TRANSMEM 692 692  
 CC TRANSMEM 693 693  
 CC TRANSMEM 694 694  
 CC TRANSMEM 695 695  
 CC TRANSMEM 696 696  
 CC TRANSMEM 697 697  
 CC TRANSMEM 698 698  
 CC TRANSMEM 699 699  
 CC TRANSMEM 700 700  
 CC TRANSMEM 701 701  
 CC TRANSMEM 702 702  
 CC TRANSMEM 703 703  
 CC TRANSMEM 704 704  
 CC TRANSMEM 705 705  
 CC TRANSMEM 706 706  
 CC TRANSMEM 707 707  
 CC TRANSMEM 708 708  
 CC TRANSMEM 709 709  
 CC TRANSMEM 710 710  
 CC TRANSMEM 711 711  
 CC TRANSMEM 712 712  
 CC TRANSMEM 713 713  
 CC TRANSMEM 714 714  
 CC TRANSMEM 715 715  
 CC TRANSMEM 716 716  
 CC TRANSMEM 717 717  
 CC TRANSMEM 718 718  
 CC TRANSMEM 719 719  
 CC TRANSMEM 720 720  
 CC TRANSMEM 721 721  
 CC TRANSMEM 722 722  
 CC TRANSMEM 723 723  
 CC TRANSMEM 724 724  
 CC TRANSMEM 725 725  
 CC TRANSMEM 726 726  
 CC TRANSMEM 727 727  
 CC TRANSMEM 728 728  
 CC TRANSMEM 729 729  
 CC TRANSMEM 730 730  
 CC TRANSMEM 731 731  
 CC TRANSMEM 732 732  
 CC TRANSMEM 733 733  
 CC TRANSMEM 734 734  
 CC TRANSMEM 735 735  
 CC TRANSMEM 736 736  
 CC TRANSMEM 737 737  
 CC TRANSMEM 738 738  
 CC TRANSMEM 739 739  
 CC TRANSMEM 740 740  
 CC TRANSMEM 741 741  
 CC TRANSMEM 742 742  
 CC TRANSMEM 743 743  
 CC TRANSMEM 744 744  
 CC TRANSMEM 745 745  
 CC TRANSMEM 746 746  
 CC TRANSMEM 747 747  
 CC TRANSMEM 748 748  
 CC TRANSMEM 749 749  
 CC TRANSMEM 750 750  
 CC TRANSMEM 751 751  
 CC TRANSMEM 752 752  
 CC TRANSMEM 753 753  
 CC TRANSMEM 754 754  
 CC TRANSMEM 755 75



Db 280 H-----TFLIGCFSLPILITRIGKKYKFLPMLAPLPLAVVSLWVFLRKAD 330  
Qy 211 PPHPI-----LSSLSI--LFFV--HOGTVKSKFSLSVTRIRPIIVM----- 248  
Db 331 -FHGVTVVRIHIGKGLNPMMSIODDFMPLHGQIAKIGLIIAIVALTBAALVGRSFAKING 389  
Qy 249 YMONALKEOGHGLSRYL--FRCCYCCFMCIDKYLHLNQNAVTTVAINGDPTCSAXDA 306  
Db 390 YLIDGKHEVVALGFMMVLGSLFYSCTAA-----TGSSTRIVN---PAACGCTA 434  
Qy 307 FYLSKNSHSFTSINCFGDPILF--LGKVLVGCFTVFGGLMAFNRYRAFOVWAVP----- 359  
Db 435 MENIVAAVVFVALBCTRLRLYYTPAIIASIIISLAPGL--NINEALHIWVKDFDEL 492  
Qy 360 LLLVAFAYLVVHSLFSLVFETVLDALFLCPA 390  
Db 493 ALIIGAFEGVL---PASVEIGLLVAVVISFA 519

```

RESULT 12
ID Y294 MYCPN STANDARD; PRT; 475 AA.
AC P75366;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG294 homolog (A05_orf475).
GN MPN421 OR MP420.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirk E., Li B.-C.,
RT Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RT Nucleic Acids Res. 24:4420-4449(1996).
RL -1 SIMILARITY: SOME, TO E.COLI YIHN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AB000041; AAB96068.1; -.
DR PIR, S73746; S73746.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 475 AA; 52731 MW; 5A1F1F0E499B86D CRC64;

```

Query Match	4.5%	Score 104.5	DB 1	Length 475
Best Local Similarity	20.0%	Pred. No. 1.3		
Matches 78; Conservative	56	Mismatches 153		Indels 103; Gaps 16

[illegible][illegible]

RESULT 13	
NPT1_RAT	
ID	NPT1_RAT
1	0.000000
	STANDARD;
	PRT; 465 AA

```

DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Renal sodium-dependent phosphate transporter protein 1 (Sodium/phosphate
DE      cotransporter 1) (Na(+)/Pi cotransporter 1) (Renal sodium-phosphate
DE      transporter protein 1) (Renal Na(+)-dependent phosphate cotransporter
DE      1) .
GN      Slc127A1 OR NPT1.
OS      Rattus norvegicus (Rat) .
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurgnath; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Liver;
RX      MEDLINE=9702:433; PubMed=8867793;
RA      Li H., Xie Z.,
RT      "Molecular cloning of two rat Na+/Pi cotransporters: evidence for
RT      differential tissue expression of transcripts.";
RL      Cell. Mol. Biol. Res. 41:451-460(1995) .
CC      -1- FUNCTION: Important for the reorption of phosphate by the kidney.
CC      May be involved in actively transporting phosphate into cells via
CC      Na(+)-cotransport in the renal brush border membrane.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -1- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC      or send an email to license@ebi.ac.uk .

```

[illegible]

FT CARBOHYD 56 56 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SEQ SEQUENCE 465 AA; 51350 MW; 288BFB8D5C0AC52 CRC64;  
 Query Match 4.4%; Score 103; DB 1; Length 465;  
 Best Local Similarity 21.1%; Pred. No. 1.6;  
 Matches 71; Conservative 39; Mismatches 95; Indels 132; Gaps 17;  
 QY 36 FLSVITLLFPVCGVLMVLDYDNDLSIEDTERENKCVLGRAIVSTGTAVLILF 95  
 DB 181 FIALVSG--FICDLGHPVFTL-----IGVGVSLFWLTLF 219  
 QY 96 -----VLRKIKLTVELFQITNKAISSAP--FLFQPLMTFALLFFWVLMVAVL 143  
 DB 220 DDPMNHFMVSSSEKQYITSSIMQVHSGRSLPIKAMLSLPLM--AII----- 266  
 QY 144 LSLGRGAAGVMEGGQVEYKPLSGIRVMSYHLGLTSTSEFLLACQMTAGAVTCTF 203  
 DB 267 -----LNSPAFIWSNNLL-VYTPPTFI-----STLHV 293  
 QY 204 NRSKNDPDHPLLSLSILFPYHQTGV--KGSFLI-----SVRIPIIWMQNALKE 256  
 DB 294 NVRENG-----LLSLPYLLAYIGIVAGMSDPLSKIRSVAVAKLFTTL----- 341  
 QY 257 QQHGALSTYLRCCYCCWCWCDKYLHINQNAVTT--TAINGT--DQTSAKDAFKI 309  
 DB 342 --GIFCPVIVVC-----LLYLSYMFSTVIFLTANSTLSFSGQLNALDI 388  
 QY 310 LSKNSHPTSINCFDPIIFLQKVLVVCFTVEGGMA 346  
 DB 389 APR-----YGFPL-KAVTALIGIFGLIS 411  
 RESULT 14  
 FLOI CRIGR STANDARD; PRT; 518 AA.  
 ID FLOI CRIGR STANDARD; PRT; 518 AA.  
 AC P42557;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Folate transporter 1 (Folate carrier protein) (Methotrexate uptake protein).  
 GN STC19A1.  
 OS Cricetus griseus (Chinese hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.  
 CC NCBI\_TaxID=10029;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RX MEDLINE=94164933; PubMed=8119923;  
 RA Williams F.M.R., Murray R.C., Underhill T.M., Flintoff W.F.; "Isolation of a hamster cDNA clone coding for a function involved in methotrexate uptake.";  
 RT J. Biol. Chem. 269:5810-5816(1994).  
 CC -1- FUNCTION: TRANSPORTER FOR THE INTAKE OF FOLATE. INVOLVED IN METHOTREXATE UPTAKE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE STC19A FAMILY OF TRANSPORTERS.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; U03031; AAC52138.1; -.  
 DR PIR; A53207; A53207.  
 DR InterPro; IPR002666; Folate carrier.  
 DR Pfam; PF01770; Folate carrier; 1.  
 DR TIGRfams; TIGR00806; ftc; 1.

KM Folate-binding; Transport; Transmembrane; Glycoprotein.  
 FT TRANSMEM 28 48 POTENTIAL.  
 FT TRANSMEM 65 85 POTENTIAL.  
 FT TRANSMEM 91 111 POTENTIAL.  
 FT TRANSMEM 122 142 POTENTIAL.  
 FT TRANSMEM 156 176 POTENTIAL.  
 FT TRANSMEM 180 202 POTENTIAL.  
 FT TRANSMEM 274 294 POTENTIAL.  
 FT TRANSMEM 304 324 POTENTIAL.  
 FT TRANSMEM 331 351 POTENTIAL.  
 FT TRANSMEM 362 382 POTENTIAL.  
 FT TRANSMEM 398 418 POTENTIAL.  
 FT TRANSMEM 432 452 POTENTIAL.  
 SQ SEQUENCE 518 AA; 58611 MW; 0B18267A134FC5A CRC64;  
 Query Match 4.4%; Score 103; DB 1; Length 518;  
 Best Local Similarity 19.3%; Pred. No. 1.8;  
 Matches 76; Conservative 68; Mismatches 133; Indels 116; Gaps 20;  
 QY 67 DTERENKCVLGRAIVSTGTAVL-----LVLIIFVLRKIKLTVELFQITNKAISSAP-- 119  
 DB 19 DRELKMRGLV-FYLCFEGMAQLRPGSEFIPYLQ--NTTIE--QVTNELLFVLYS 73  
 QY 120 -----FLFQPLMTFALLFFWVLMVAVLISLGTAGA-AQVMEGGQVEYKPL 165  
 DB 74 HLAIVPIPLTDVLRKPKILLIQLCSFMCV-W--LLILLGTSVVMQDME--VFYSVT 127  
 QY 166 SGIRVMSYHLGLTSTSEF-----ILACQMTAGAVTCTFENSKNDPDHP 214  
 DB 128 MARIATYSYISLVRPSRYOMASYRAAVLIGVTSSTVAGQVIMPEQSSONS----- 182  
 QY 215 ILSLSILFPYHQTGVKGSFLISVRIPIIWMQNAL-----KEQOH----- 259  
 DB 183 -----NMLNYISLGIIFSLGLSLPKRPSLFFNRSAVYKALPCBLDQHPGAPBE 237  
 QY 260 -GALSRYLPRC-----CYCCWCWCDK-----YLH-----LN 285  
 DB 238 PGLTELVGSGNSFLVCLSELVGNLRQPHVRLCWMVRFASAGYIIVYVHVLMSID 297  
 QY 286 QNAVYTTAINGTDFCTSAKDAFKILSKNSHPTSINCFDPIIFLQKVLVVCFTVEGGIM 345  
 DB 298 KVLNNGAVDASTLSLTSF-----SAGVYKIR-----WALMSKVLVASVIALQAGLV 347  
 QY 346 AATYNAFQVAVPPLLVAFFAYLV--AHSTL 375  
 DB 348 FCMYWHVYVT-VHKIWLVMYTVLFRGAYQFL 379  
 RESULT 15  
 CTA\_BACSU  
 ID CTA\_BACSU STANDARD; PRT; 306 AA.  
 AC P12946;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cytochrome A33 controlling protein.  
 GN CTA OR BSU14870.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CC NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=89359135; PubMed=2549006;  
 RA Mueller J.P., Taber H.W.; "Isolation and sequence of ctaA, a gene required for cytochrome a33 biosynthesis and sporulation in Bacillus subtilis.";  
 RT J. Bacteriol. 171:4967-4978(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Bertero M., Presciani E., Glaser P., Richou A., Danchin A.; "Bacillus subtilis chromosomal region downstream nprE.";  
 RT Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.

[3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertolo M.G., Bessières P., Bolotin A., Borchert S.,  
 RA Borries R., Bourlès L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Brusch C.V., Caldwell B., Capiano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Comerfort I.F., Cummings N.J., Daniel R.A.,  
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Ertien K.D., Errington J., Fabret C., Ferrati E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
 RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Prescan E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadala Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Serot S.J., Serot P., Shin B.S., Soldo B.,  
 RA Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretto A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,  
 RA Winters P., Wipart A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.,  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256 (1997).  
 CC -!- FUNCTION: CTA is REQUIRED FOR CYTOCHROME AA3 BIOSYNTHESIS AND  
 CC INDEPENDENTLY FOR SPOULATION  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL, M23915; AAA50254.1; -  
 DR EMBL, Z98682; CAB11340.1; -  
 DR EMBL, Z99111; CAB13360.1; -  
 DR PIR, A33960; A33960.  
 DR Subtilist; BG10213; ctaa.  
 DR InterPro; IPR003780; COX15\_Ctaa.  
 DR Pfam; PF02628; COX15\_Ctaa; 1.  
 KW Sporulation; Membrane; Complete proteome.  
 SQ SEQUENCE 306 AA; 34085 MW; D47FE9041851CDDC CRC64;

Query Match 4.4%; Score 102.5; DB 1; Length 306;  
 Best Local Similarity 21.7%; Pred. No. 1.2; Mismatches 53; Gaps 10;  
 Matches 53; Conservative 39; Indels 87; Gaps 65; Indels 10;

QY 9 GUCIALALSLAMFTFRTITLVH-IFISLVIGLIFVCGVLMWLYDYNDISIELD 67  
 DB 65 GISII-LVYSLA-FMSWKITPIPRETPLAISIIPIFLQALTL-----GALAAVFG 114  
 QY 68 TERENMKCVLGFATVSTGTAVLVLIPLAKRIKLYELFPI----- 110  
 DB 115 SNALIMALHFGISLISFASVLLITLILFEADKSVRLVPLQIGKMOFHMIGILIVSYI 174  
 QY 111 -----TKKAISAPFLFOPV-----W-----TPAILIFFWLVAVAVL 144  
 DB 175 VVYTGAYVHTSSSLACPNVPLCSPLNNGLPQGFHEWQMGHRAAALLLFWIITVAAVHA 234

QY 145 SLGTAGAQQVMEGG-----QVEYKPLSGIRYMSYHLISLIMTSEFLACQQTITGAVY 159  
 DB 235 ITSYKQOKQIFWGWISCLIFITTLQALSGIMIVISELAGFALAHSPFIAC-----LFG--V 288  
 QY 200 TCYP 203  
 DB 289 LCVF 292  
 Search completed: May 27, 2004, 16:36:02  
 Job time : 20 secs

Tue Jun 1 09:37:58 2004

us-10-063-686-32.ra1

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:34:25 ; Search time 22 Seconds  
(without alignments)  
1044.253 Million cell updates/sec

Title: US-10-063-686-32  
Perfect score: 2315  
Sequence: 1 MSGRDTLGLCTIALALISLA.....QDDKSLRNEGTLEQATVR 445

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1aa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207	8.9	116	US-08-702-344-21	Sequence 21, Appl
2	161.5	7.0	148	US-09-247-155-166	Sequence 166, Appl
3	112.5	4.9	348	US-08-513-974B-46	Sequence 46, Appl
4	112.5	4.9	348	US-08-993-088A-10	Sequence 10, Appl
5	112.5	4.9	348	US-08-993-424B-10	Sequence 10, Appl
6	112.5	4.9	348	US-08-540-650B-2	Sequence 2, Appl
7	112.5	4.9	348	US-09-595-549-5	Sequence 5, Appl
8	112.5	4.9	348	US-09-461-436B-46	Sequence 46, Appl
9	112.5	4.9	348	US-09-603-680-10	Sequence 10, Appl
10	110.5	4.8	348	US-08-513-974B-342	Sequence 342, Appl
11	108.5	4.7	346	US-09-199-737-5	Sequence 5, Appl
12	108.5	4.7	346	US-08-993-088A-3	Sequence 3, Appl
13	108.5	4.7	346	US-08-993-424B-3	Sequence 3, Appl
14	108.5	4.7	346	US-09-058-333A-5	Sequence 5, Appl
15	108.5	4.7	346	US-09-595-549-6	Sequence 6, Appl
16	108.5	4.7	346	US-09-603-680-3	Sequence 3, Appl
17	107.5	4.7	346	US-08-899-112B-30	Sequence 30, Appl
18	107.5	4.6	349	US-08-513-974B-343	Sequence 343, Appl
19	107	4.6	345	US-08-981-700A-5	Sequence 5, Appl
20	106	4.6	497	US-09-718-693A-1	Sequence 1, Appl
21	102.5	4.4	395	US-08-900-230-5	Sequence 5, Appl
22	102	4.4	378	US-09-045-583-5	Sequence 5, Appl
23	102	4.4	378	US-09-534-185-5	Sequence 5, Appl
24	102	4.4	394	US-09-134-000C-3732	Sequence 3732, Ap
25	102	4.4	446	US-07-626-618A-21	Sequence 21, Appl
26	102	4.4	446	US-08-333-977-21	Sequence 21, Appl
27	102	4.4	446	US-07-969-267B-4	Sequence 4, Appl

28	102	4.4	446	US-09-168-510-4	Sequence 4, Appl
29	102	4.4	606	US-09-107-532A-4683	Sequence 4683, Ap
30	99.5	4.3	597	US-09-328-352-4703	Sequence 4703, Ap
31	99	4.3	940	US-09-328-352-8165	Sequence 8165, Ap
32	98.5	4.3	889	US-08-118-101A-4	Sequence 4, Appl
33	98	4.2	318	US-08-345-696-25	Sequence 25, Appl
34	98	4.2	318	US-08-233-009-25	Sequence 25, Appl
35	98	4.2	318	US-08-560-231-25	Sequence 25, Appl
36	98	4.2	318	US-09-080-704A-25	Sequence 25, Appl
37	98	4.2	318	US-09-800-274-6	Sequence 6, Appl
38	98	4.2	486	US-09-134-001C-3593	Sequence 3593, Ap
39	98	4.2	663	US-09-711-164-352	Sequence 352, Ap
40	97.5	4.2	374	US-09-102-710B-3	Sequence 3, Appl
41	97	4.2	335	US-09-387-699-4	Sequence 4, Appl
42	97	4.2	335	US-09-641-259B-4	Sequence 4, Appl
43	97	4.2	454	US-09-252-991A-30013	Sequence 30013, A
44	97	4.2	613	US-09-170-486D-212	Sequence 212, App
45	96.5	4.2	454	US-09-134-001C-3501	Sequence 3501, Ap

#### ALIGNMENTS

RESULT 1  
US-08-702-344-21  
Sequence 21, Application US/08702344

Patent No. 5723315  
GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John  
APPLICANT: Lavallie, Edward  
APPLICANT: Racie, Lisa  
APPLICANT: Treacy, David  
APPLICANT: Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
TITLE OF INVENTION: ENCODING THEM  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,344  
FILING DATE:

CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS:

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-702-344-21

Query Match 8.9%; Score 207; DB 1; Length 116;  
Best Local Similarity 38.3%; Pred. No. 2e+14;  
Matches 46; Conservative 19; Mismatches 29; Indels 26; Gaps 4;

Tue Jun 1 09:37:58 2004

us-10-063-686-32.ra1

Page 2

QY	269	CCGCCGACGDKYKLLHNNQAYTTTANGNDPCTSNACD	EKLISKNSHSPTISINCGPFI	32 8
		:::     :::     :::     :::     :::		
Db	3	CLMGCFMCJEKFKFLPNRAVIMAIYGINFCTSRANF	FLIMRNIIIRAVLNDKVDPELF	62
QY	329	FLKRLVYCTVVGGLMAF-----	NTNRAQWNAVPLLVAFAFLVA	371
		:::     :::     :::     :::     :::		
Db	63	ILGKLLIVSVY---GLIAFFFTTHRIYQDAPLAN---	YV-PELILVIVGSLIA	113

```

RESULT 2
US-09-247-155-166
Sequence 166, Application US/09247155A
Patent No. 6312922
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclert, Americ
APPLICANT: Bouguetelert, Lydie
TITLE OF INVENTION: Complementary DNAs
FILE REFERENCE: GENSET.021A
CURRENT APPLICATION NUMBER: US/09/247, 155A
CURRENT FILING DATE: 1999-02-09
EARLIER APPLICATION NUMBER: 60/074,121
EARLIER FILING DATE: 1998-02-09
EARLIER APPLICATION NUMBER: 60/081,563
EARLIER FILING DATE: 1998-04-13
EARLIER APPLICATION NUMBER: 60/096,116
EARLIER FILING DATE: 1998-08-10
EARLIER APPLICATION NUMBER: 60/099,273
EARLIER FILING DATE: 1998-10-04
NUMBER OF SEQ ID NOS: 182
SOFTWARE: Patent.pm
SEQ ID NO 166
LENGTH: 148
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -48...-1
FEATURE:
NAME/KEY: UNSTRE
LOCATION: 32,100
OTHER INFORMATION: Xaa = any one of the twenty amino acids
US-09-247-155-166

```

[illegible]

RESULT 3  
 US-08-513-974B-46  
 ; Sequence 46, Application US/08513574B  
 ; Patent No. 6114139  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hinuma, Shuji  
 ; APPLICANT: Hosoya, Masaki  
 ; APPLICANT: Fujii, Ryo  
 ; APPLICANT: Ohtaki, Tetuya  
 ; APPLICANT: Fukusumi, Shoji  
 ; APPLICANT: Ongi, Kazuhito  
 TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN

```

1 TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
2 NUMBER OF SEQUENCES: 380
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
5 STREET: 130 Water Street
6 CITY: Boston
7 STATE: MA
8 COUNTRY: USA
9 ZIP: 02109
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Patentin Release #1.0, Version #1.30
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/513,974B
18 FILING DATE: 14-SEP-1995
19 CLASSIFICATION: 536
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: PCT/JP95/01599
22 FILING DATE: 10-AUG-1995
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: JP 7-093989
25 FILING DATE: 19-AUG-1995
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: JP 7-057186
28 FILING DATE: 16-MAR-1995
29 PRIOR APPLICATION DATA:
30 APPLICATION NUMBER: JP 7-007177
31 FILING DATE: 20-JAN-1995
32 PRIOR APPLICATION DATA:
33 APPLICATION NUMBER: JP 6-36611
34 FILING DATE: 28-DEC-1994
35 PRIOR APPLICATION DATA:
36 APPLICATION NUMBER: JP 6-270017
37 FILING DATE: 02-NOV-1994
38 PRIOR APPLICATION DATA:
39 APPLICATION NUMBER: JP 6-236357
40 FILING DATE: 30-SEP-1994
41 PRIOR APPLICATION DATA:
42 APPLICATION NUMBER: JP 6-236356
43 FILING DATE: 30-SEP-1994
44 PRIOR APPLICATION DATA:
45 APPLICATION NUMBER: JP 6-189274
46 FILING DATE: 11-AUG-1994
47 PRIOR APPLICATION DATA:
48 APPLICATION NUMBER: JP 6-189273
49 FILING DATE: 11-AUG-1994
50 PRIOR APPLICATION DATA:
51 APPLICATION NUMBER: JP 6-189272
52 FILING DATE: 11-AUG-1994
53 ATTORNEY/AGENT INFORMATION:
54 NAME: Reenick, David S.
55 REGISTRATION NUMBER: 34,235
56 REFERENCE/DOCKET INFORMATION: 45753
57 TELECOMMUNICATION INFORMATION:
58 TELEPHONE: 617-523-3400
59 TELEFAX: 617-523-6440
60 INFORMATION FOR SEQ ID NO: 46:
61 SEQUENCE CHARACTERISTICS:
62 LENGTH: 348 amino acids
63 TYPE: amino acid
64 STRANDEDNESS:
65 TOPOLOGY: linear
66 MOLECULE TYPE: peptide
67
68 US-08-513-974B-46

```

Query Match	4.9%;	Score 112.5;	DB 3;	Length 348;
Best Local Similarity	19.4%;	Pred. No. 0.0013;		
Matches	83;	Conservative 43;	Mismatches 129;	Indels 173;
			Gaps 177;	

33 VHFISLVILGLHFFVCGVLMVLYDYDNDLSIEDLTERENMKCVLGFAIVSTGIRAVLVV 92

```

Db      31 VENTITLVFGLIFAMGVL-----GNSLVITV 57
Qy      93 LIFVLRRIKLTVELFOITNKAISSAPLLF-----OPLMTFAILL---FWVLW 139
Db      58 LARSKPGKPRSTYLF--ILNLSIADLAVLLFCIPQATVVALPTWVGAFICKFIHYFF 116
Qy      140 VAVLLSCTAGAAQV-----MEGQVEYKPLSGIRYMSYHLIGLMTSEFLA 188
Db      117 VSMVLSITFLAASVDRYVAIVHSRRSSLRVSNALLGVGFIVA-----161
Qy      189 COOMTGAAGVTCY---FNRSKNDP-----PDHPIISSLSILFFVHGTGVKGSFLIS 238
Db      162 --LSIMASBPVAVHQRLFHRDSNQTFCMEQWPNK-----LHKAYVVCVFVG 207
Qy      239 VVRIPIRIWMONALKKEQHGALSRYLFRCCYCCFWCLDKYLLHLNQNATTTAINGTD 298
Db      208 -----YLLPILLICF-CYAKVNLHLHK-----229
Qy      299 FCTSAKDAFKILSNSSHFTSINCFGDFIIFLGKLVVCFVFGMAFNRAFOVW--356
Db      230 -----LKNMSKSEASKKT-----AQTVLVVVVFG--ISMLPHVHLMAB 270
Qy      357 --AVPLLVAFFAYLVAH-----SFLSV-FETVLDALFLCPAVDLETNDGS 399
Db      271 FGAFPLTPASFFRITAHCLAVSNNSVNPITVAFLSNFRKAYKQVFKCHVCDESPRSET 330
Qy      400 SEKPYFMD 407
Db      331 KENKSRMD 338

```

## RESULT 4

```

US-08-993-088A-10
; Sequence 10, Application US/08993088A
; Patent No. 6287855
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Sullivan, Kathleen
; TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,088A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 19846
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1958
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-993-088A-10

Query Match 4.9%; Score 112.5; DB 3; Length 348;  
 Best Local Similarity 19.4%; Pred. No. 0.0013;  
 Matches 83; Conservative 43; Mismatches 129; Indels 173; Gaps 17;

```

Qy      33 VHIISLVILGLIFVCGVLMVLYDYDTNLSIELDTERENKCVLGFALVSTGITAVLV 92
Db      31 VENTITLVFGLIFAMGVL-----GNSLVITV 57
Qy      93 LIFVLRRIKLTVELFOITNKAISSAPLLF-----OPLMTFAILL---FWVLW 139
Db      58 LARSKPGKPRSTYLF--ILNLSIADLAVLLFCIPQATVVALPTWVGAFICKFIHYFF 116
Qy      140 VAVLLSCTAGAAQV-----MEGQVEYKPLSGIRYMSYHLIGLMTSEFLA 188
Db      117 VSMVLSITFLAASVDRYVAIVHSRRSSLRVSNALLGVGFIVA-----161
Qy      189 COOMTGAAGVTCY---FNRSKNDP-----PDHPIISSLSILFFVHGTGVKGSFLIS 238
Db      162 --LSIMASBPVAVHQRLFHRDSNQTFCMEQWPNK-----LHKAYVVCVFVG 207
Qy      239 VVRIPIRIWMONALKKEQHGALSRYLFRCCYCCFWCLDKYLLHLNQNATTTAINGTD 298
Db      208 -----YLLPILLICF-CYAKVNLHLHK-----229
Qy      299 FCTSAKDAFKILSNSSHFTSINCFGDFIIFLGKLVVCFVFGMAFNRAFOVW--356
Db      230 -----LKNMSKSEASKKT-----AQTVLVVVVFG--ISMLPHVHLMAB 270
Qy      357 --AVPLLVAFFAYLVAH-----SFLSV-FETVLDALFLCPAVDLETNDGS 399
Db      271 FGAFPLTPASFFRITAHCLAVSNNSVNPITVAFLSNFRKAYKQVFKCHVCDESPRSET 330
Qy      400 SEKPYFMD 407
Db      331 KENKSRMD 338

```

## RESULT 5

```

US-08-993-424B-10
; Sequence 10, Application US/08993424B
; Patent No. 6337206
; GENERAL INFORMATION:
; APPLICANT: Tan, Carina
; APPLICANT: Koliakowski, Lee F., Jr.
; TITLE OF INVENTION: MOUSE GALANIN RECEPTOR GALR2 AND
; TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,424B
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,851
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179

```

Tue Jun 1 09:37:58 2004

us-10-063-686-32.fai

Page 4

REFERENCE/DOCKET NUMBER: 19846NP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1958  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-993-424B-10

Query Match 4.9%; Score 112.5; DB 4; Length 348;  
Best Local Similarity 19.4%; Pred. No. 0.0013;  
Matches 83; Conservative 43; Mismatches 129; Indels 173; Gaps 17;

QY 33 VHFISLVILGLIFVCGVLMVLYDYTDNLSIEDTERENMKCVLGFALVSTGTTAVLLV 92  
DB 31 VENFTLVVFGILFAMGVL-----GNSLVITV 57  
QY 93 LIFVLRKRRIKLTVELFOITNKAISSAPFLF-----OPLMTFAILI--FFVWLW 139  
DB 58 LARSKGKPRSTTNLF--ILNLSIADLAVILFCIPQATVVALPTVWLGAFICKFIHYEFT 116  
QY 140 VAVILSLGTAGAAQV-----MEGGQVEYKPLSGIRYMSYHLIGLIWTSFILA 188  
DB 117 VSMIVSIFTLAAMSDRYVALVHSRRSSSLRVSRNALGVGPIWA-----161  
QY 189 CQOMTIAGAVVTCY---FNRSKNDP-----PDHPIISSLSILFFYHOGTVVKGSPILS 238  
DB 162 ---LSIAMSFPVAYHQRLFHRDSNQTFCWEQWPNK-----LHKKAYVCTTFVFG 207  
QY 239 VVRIRIIVWMQNALKEQHGALSRYLPRCCYCCFWCLDKYLHLNQNATTTAINGTD 298  
DB 208 -----YLLPLLLICF-CYAKVLNHLHKK-----229  
QY 239 FCTSAKDAPKILSKNSHFTSINCFDPIIFLGKVLVCFVFGGLMAFNVARAFOVW-- 356  
DB 230 -----LKNMSKSEASKKT-----AQTVLVVVVFG--ISWLPHHVHLMAE 270  
QY 357 --AVPLLVAFPAIVAH-----SPLSV-FETVLDALFLCFVADLETNDGS 399  
DB 271 FGAFLTPASFFPRITAHCLAVSNSVNPPIYAFLENPRAYKQVFKCHVCDSPRSET 330  
QY 400 SEKPYFMD 407  
DB 331 KENKSRMD 338

RESULT 6  
US-08-540-650B-2  
Sequence 2, Application US/08540650B  
Patent No. 6399325  
GENERAL INFORMATION:  
APPLICANT: HINUMA, Shuji  
APPLICANT: FUJII, Ryo  
APPLICANT: FUKUSUMI, Shoji  
APPLICANT: OHYAKI, Tetsuya  
APPLICANT: HOSOGA, Masaki  
APPLICANT: OHGI, Kazuhiro  
APPLICANT: ONDA, Haruo  
TITLE OF INVENTION: GALANIN RECEPTOR PROTEIN, PRODUCTION AND USE THEREOF  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN  
STREET: 130 WATER STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: US  
ZIP: 02109  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/540,650B  
FILING DATE: 11-OCT-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 7-134412  
FILING DATE: 31-MAY-1995  
APPLICATION NUMBER: 6-326610  
FILING DATE: 28-DEC-1994  
APPLICATION NUMBER: 6-247599  
FILING DATE: 13-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: RESNICK, DAVID S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45901  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348  
TYPE: Amino acid  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-540-650B-2

Query Match 4.9%; Score 112.5; DB 4; Length 348;  
Best Local Similarity 19.4%; Pred. No. 0.0013;  
Matches 83; Conservative 43; Mismatches 129; Indels 173; Gaps 17;

QY 33 VHFISLVILGLIFVCGVLMVLYDYTDNLSIEDTERENMKCVLGFALVSTGTTAVLLV 92  
DB 31 VENFTLVVFGILFAMGVL-----GNSLVITV 57  
QY 93 LIFVLRKRRIKLTVELFOITNKAISSAPFLF-----OPLMTFAILI--FFVWLW 139  
DB 58 LARSKGKPRSTTNLF--ILNLSIADLAVILFCIPQATVVALPTVWLGAFICKFIHYEFT 116  
QY 140 VAVILSLGTAGAAQV-----MEGGQVEYKPLSGIRYMSYHLIGLIWTSFILA 188  
DB 117 VSMIVSIFTLAAMSDRYVALVHSRRSSSLRVSRNALGVGPIWA-----161  
QY 189 CQOMTIAGAVVTCY---FNRSKNDP-----PDHPIISSLSILFFYHOGTVVKGSPILS 238  
DB 162 ---LSIAMSFPVAYHQRLFHRDSNQTFCWEQWPNK-----LHKKAYVCTTFVFG 207  
QY 239 VVRIRIIVWMQNALKEQHGALSRYLPRCCYCCFWCLDKYLHLNQNATTTAINGTD 298  
DB 208 -----YLLPLLLICF-CYAKVLNHLHKK-----229  
QY 239 FCTSAKDAPKILSKNSHFTSINCFDPIIFLGKVLVCFVFGGLMAFNVARAFOVW-- 356  
DB 230 -----LKNMSKSEASKKT-----AQTVLVVVVFG--ISWLPHHVHLMAE 270  
QY 357 --AVPLLVAFPAIVAH-----SPLSV-FETVLDALFLCFVADLETNDGS 399  
DB 271 FGAFLTPASFFPRITAHCLAVSNSVNPPIYAFLENPRAYKQVFKCHVCDSPRSET 330  
QY 400 SEKPYFMD 407  
DB 331 KENKSRMD 338

RESULT 7  
US-09-595-549-5  
Sequence 5, Application US/09595549  
Patent No. 651827  
GENERAL INFORMATION:  
APPLICANT: Howard, Andrew D.

APPLICANT: Cascieri, Margaret A.  
APPLICANT: Smith, Roy G.  
APPLICANT: Sullivan, Kathleen A.  
APPLICANT: Tan, Carina  
APPLICANT: Van der Ploeg, Leonardus H. T.  
APPLICANT: Lynch, Kevin R.  
TITLE OF INVENTION: GALANIN RECEPTOR GALT3 AND NUCLEOTIDES  
FILE REFERENCE: 20148PCA  
CURRENT APPLICATION NUMBER: US/09/595,549  
PRIOR FILING DATE: 2000-06-16  
PRIOR APPLICATION NUMBER: US98/26812  
PRIOR FILING DATE: 1998-12-17  
PRIOR APPLICATION NUMBER: 60/069,725  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 348  
TYPE: PRT  
ORGANISM: mouse  
US-09-595-549-5

Query Match 4.9%; Score 112.5; DB 4; Length 348;  
Best Local Similarity 19.4%; Pred. No. 0.0013;  
Matches 83; Conservative 43; Mismatches 129; Indels 173; Gaps 17;

QY 33 VHIISLVILGLFVCGVLMWLYDYNDISIELDERENMKCVLGFALVSTGTAVALV 92  
DB 31 VENFTLVVGLIFAMGVL-----GNSLVITV 57  
QY 93 LIFVLRKRIKLTVELFOITNKATISAPFLF-----QPLMTFALLI---FFWVLM 139  
DB 58 LARSKPKRSTNLF-ILNLSIADLAVLFCIPQATYVALFTWVGAFICKFIHYFFT 116  
QY 140 VAVLSLGTAGAAY-----MEGQVEYKPLSGIRYMSYHLGLIWTSEFLA 188  
DB 117 VSMVLVSIPTLAASVDRYVAIVHSRSSLRVSRNALGVGFMA-----161  
QY 189 CQOMTIGAVTVC---FNRKNDP-----PDHPILSLSILFFHQGTVYKGSFLIS 238  
DB 162 ---LSIAVMSPAVYHQRLFHRSNQTFCWEQWPNK-----LHKAVVCTFVFG 207  
QY 239 VVRIPRIIVYQNALKEQOQHALSRVFCYCCFCWCDKXILHLNQAVYTTAINGTD 298  
DB 208 -----YLLPDLICF-CYAKYLNHLKX-----229  
QY 299 FCTSAADAFKILSKNSHFTSINCFGDFIIFGKVLVVCFTVFGMAFNVRAPQVW-- 356  
DB 230 -----LKNMSKXSEASKKT-----AQTVLVVVVVG-ISWLPFHVHLWAE 270  
QY 357 --AVPLVAFPAVLAH-----SPLSV-FETVLDAFLCEFADLTNGNS 399  
DB 271 FGAFFLTPASFPRRIYHCLAVSNSSVNPITYAFISENFRKAYKQVFKCHVCDSPRSET 330  
QY 400 SEKPYFMD 407  
DB 331 KENKSRMD 338

RESULT 8  
US-09-461-436B-46  
Sequence 46, Application US/09461436B  
Patent No. 6538107  
GENERAL INFORMATION:  
APPLICANT: Shuji Hinuma  
Yasunaki Ito  
Ryo Fujii  
TITLE OF INVENTION: G Protein Coupled Receptor Protein,  
Production, And Use Thereof  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edwards & Angel, LLP

STREET: 101 Federal Street  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02209  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/461,436B  
FILING DATE: 14-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/513,974  
FILING DATE: 14-SEP-1995  
APPLICATION NUMBER: PCT/JP95/01599  
FILING DATE: 10-AUG-1995  
APPLICATION NUMBER: 7-093989  
FILING DATE: 19-APR-1995  
APPLICATION NUMBER: 7-057186  
FILING DATE: 16-MAR-1995  
APPLICATION NUMBER: 7-007177  
FILING DATE: 20-JAN-1995  
APPLICATION NUMBER: 6-326611  
FILING DATE: 28-DEC-1994  
APPLICATION NUMBER: 6-270017  
FILING DATE: 02-NOV-1994  
APPLICATION NUMBER: 6-236357  
FILING DATE: 30-SEP-1994  
APPLICATION NUMBER: 6-236356  
FILING DATE: 30-SEP-1994  
APPLICATION NUMBER: 6-189274  
FILING DATE: 11-AUG-1994  
APPLICATION NUMBER: 6-189273  
FILING DATE: 11-AUG-1994  
APPLICATION NUMBER: 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CONLIN, DAVID G.  
REGISTRATION NUMBER: <Unknown>  
REFERENCE/DOCKET NUMBER: 45753 DIV2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-439-4444  
TELEFAX: 617-439-4170  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 46:  
US-09-461-436B-46

Query Match 4.9%; Score 112.5; DB 4; Length 348;  
Best Local Similarity 19.4%; Pred. No. 0.0013;  
Matches 83; Conservative 43; Mismatches 129; Indels 173; Gaps 17;

QY 33 VHIISLVILGLFVCGVLMWLYDYNDISIELDERENMKCVLGFALVSTGTAVALV 92  
DB 31 VENFTLVVGLIFAMGVL-----GNSLVITV 57  
QY 93 LIFVLRKRIKLTVELFOITNKATISAPFLF-----QPLMTFALLI---FFWVLM 139  
DB 58 LARSKPKRSTNLF-ILNLSIADLAVLFCIPQATYVALFTWVGAFICKFIHYFFT 116  
QY 140 VAVLSLGTAGAAY-----MEGQVEYKPLSGIRYMSYHLGLIWTSEFLA 188  
DB 117 VSMVLVSIPTLAASVDRYVAIVHSRSSLRVSRNALGVGFMA-----161  
QY 189 CQOMTIGAVTVC---FNRKNDP-----PDHPILSLSILFFHQGTVYKGSFLIS 238





APPLICATION NUMBER: JP 7-007177  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 342:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 348 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-513-974B-342

Query Match 4.8%; Score 110.5; DB 3; Length 348;

Best Local Similarity 19.4%; Pred. No. 0.0022;  
Matches 83; Conservative 42; Mismatches 130; Indels 173; Gaps 17;

33 VHFISVILGLFVCGVLMWLYDYDNDLSIEDTEREMKCVLGFAIVSTGTAVLLV 92  
31 VENFILLVFGILFAMGVL-----GNSLVITV 57  
93 LIFVLRKRIKLVLEFQITNKAISSAPFLF-----OPLMTFAILL--FFWLW 139  
58 LAKSKPGKPRSTNLF--ILNLSIADLAYLLFCIPQATVVALPTWVGAFICKFIHYFT 116  
140 VAVLSLGTAGAAY-----MEGQVEKPLSGIRYKMSYHLGLMTSEFLA 188  
117 VSMVLSIFTLAASVDRYVAIVHSRRSSSLRVSNALLGVGFWA----- 161  
189 COQMTIAGAVTCY-----FNRSKNDP-----PDHPISSLSILFFYHOGTVVKSFLIS 238  
162 ---LSIAMSPPVAHYQRLFHRDNGQFCWEQWPNK-----LHKAVVVCFTVFG 207  
239 VRIPIRIIWMQNALKEQHGALSRVLRCCYCCGWCCLKLHLNQNAYTTTAINGTD 298  
208 -----YLLPILLICF-CYAKVNLHLKK----- 229  
299 FCTSAKDAFKLSKNSHFTSINCFDFTIFLGKLVVCFVGGMAFNYNRAFOVW-- 356  
230 -----LKNMSKKSSEASKKT-----AQTVLVVVVFG--ISWLPHHVYHMAE 270  
357 --AVPLLVAFFAYLVAN-----SFLSV-FTVYDALFLCPAVDLETNDGS 399  
271 FGAFPLTPASFPRITAHCLAYSNSVNPITVAFLSNFRAKAYKQVFKCVCDSPRSET 330  
400 SEKPYEMD 407  
331 KENKSRMD 338

RESULT 11  
US-09-199-737-5  
Sequence 5, Application US/09199737A  
Patent No. 6287788  
GENERAL INFORMATION:  
APPLICANT: Bard, Jonathan A.  
APPLICANT: Borowsky, Beth  
APPLICANT: Smith, Kelli E.  
APPLICANT: Branchek, Theresa A.  
APPLICANT: Gerald, Christophe P.G.  
TITLE OF INVENTION: DNA Encoding Galanin GALR3 Receptors And Uses Thereof  
FILE REFERENCE: 52241-D-PCT-US  
CURRENT FILING DATE: 1998-11-25  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.0 - beta  
SEQ ID NO 5  
LENGTH: 346  
TYPE: PRT  
ORGANISM: Rat  
US-09-199-737-5

Query Match 4.7%; Score 108.5; DB 3; Length 346;  
Best Local Similarity 19.1%; Pred. No. 0.0036;  
Matches 81; Conservative 45; Mismatches 124; Indels 173; Gaps 17;

33 VHFISVILGLFVCGVLMWLYDYDNDLSIEDTEREMKCVLGFAIVSTGTAVLLV 92  
30 VENFILLVFGILFAMGVL-----GNSLVITV 56  
93 LIFVLRKRIKLVLEFQITNKAISSAPFLF-----OPLMTFAILL--FFWLW 139  
57 LAKSKPGKPRSTNLF--ILNLSIADLAYLLFCIPQATVVALPTWVGAFICKFIHYFT 115  
140 VAVLSLGTAGAAY-----MEGQVEKPLSGIRYKMSYHLGLMTSEFLA 188  
116 VSMVLSIFTLAASVDRYVAIVHSRRSSSLRVSNALLGVGFWA----- 160  
189 COQMTIAGAVTCY-----FNRSKNDP-----PDHPISSLSILFFYHOGTVVKSFLIS 238  
161 ---LSIAMSPPVAHYQRLFHRDNGQFCWEHWPNG-----LHKAVVVCFTVFG 206  
239 VRIPIRIIWMQNALKEQHGALSRVLRCCYCCGWCCLKLHLNQNAYTTTAINGTD 298  
207 -----YLLPILLICF-CYAKVNLHLKK----- 228  
299 FCTSAKDAFKLSKNSHFTSINCFDFTIFLGKLVVCFVGGMAFNYNRAFOVW-- 356  
229 -----LKNMSKKSSEASKKT-----AQTVLVVVVFG--ISWLPHHVYHMAE 269  
357 --AVPLLVAFFAYLVAN-----SFLSV-FTVYDALFLCPAVDLETNDGS 399  
270 FGAFPLTPASFPRITAHCLAYSNSVNPITVAFLSNFRAKAYKQVFKCVCDSPRSET 329  
400 SEK 402  
330 KEK 332

RESULT 12  
US-08-993-088A-3  
Sequence 3, Application US/08993088A  
Patent No. 6287855  
GENERAL INFORMATION:  
APPLICANT: Tan, Carina  
APPLICANT: Sullivan, Kathleen  
TITLE OF INVENTION: GALANIN RECEPTOR GALR2 AND  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:

Tue Jun 1 09:37:58 2004

us-10-063-686-32.fai

Page 8

ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/993,088A  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/033,651  
FILING DATE: 27-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Heber, Sheldon O.  
REGISTRATION NUMBER: 38,179  
REFERENCE/DOCKET NUMBER: 19846  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1958  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-993-088A-3

Query Match 4.7%; Score 108.5; DB 3; Length 346;  
Best Local Similarity 19.1%; Pred. No. 0.0036;  
Matches 81; Conservative 45; Mismatches 124; Indels 173; Gaps 17;

33 WHIFSLVILGLFVCGVLMVLYDYNDLSIEDTERENKCVLGPAIVSTGITAVLLV 92  
30 VENFTLVVFGILFAMGVLT-----GNSLVITV 56  
93 LIFVLKRIKLTVELFOITNKALISSAPFLF-----QPLMTFAILL---FFVWLM 139  
57 LARSKRGKRGSTTNLF--ILNLSIDLAVLLFCIPPOATVVALFTVVLGAFICKFIHYFT 115  
140 VAVILSLGTAGAAQV-----MEGQVEYKPLSGIRYMSYHLIGLIWTSEFLA 188  
116 VSMVLSIFTLAMSVDRVVAIVHSRRSSSLRVSRNALGVGFMA----- 160  
189 CQOMTIAAVTCY---FNRSKNDP-----PDHPILSSLSILFFYHOGTVVKGSPFLIS 238  
161 ---LSIAASPAVAYQRLFRHDSNQTFCWEHPNQ-----LHKAYVVCTEVFG 206  
239 VVRIPRIIVMOMALKEQOHGALSRYLFRCCYCCFWCLDKYLLHNLONAVTTAINGTD 298  
207 -----YLLPILLICF-CYAKVNLHLHK----- 228  
299 FCTSAKDAFKILSKNSHFTSINCRGDFIIFLGKVLVVCFTVPGGLMAFNVRAPQVW-- 356  
229 -----LKNMSKSEASKKT-----AQTIVLVVVVFG--ISMLEPHVHIMAE 269  
357 ---AVPLILVAFAVILVA--SFLSV-FETVLDAFLFCFAVDELFTDGS 399  
270 FGAFPLTPASFFPRTACLAVNSNSVNPILYALISENFRAYKQVKEKRCVNSPFGDA 329  
400 SEK 402  
330 KEK 332  
Db

RESULT 13

US-08-993-424B-3  
Sequence 3, Application US/08993424B  
Patent No. 6337206  
GENERAL INFORMATION:  
APPLICANT: Tan, Carina  
APPLICANT: Kolakowski, Lee F., Jr.  
TITLE OF INVENTION: MOUSE GALANIN RECEPTOR GALP2 AND  
TITLE OF INVENTION: NUCLEOTIDES ENCODING SAME  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/993,424B  
FILING DATE: 18-DEC-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/033,651  
FILING DATE: 27-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Heber, Sheldon O.  
REGISTRATION NUMBER: 38,179  
REFERENCE/DOCKET NUMBER: 19846NP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1958  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-993-424B-3

Query Match 4.7%; Score 108.5; DB 4; Length 346;  
Best Local Similarity 19.1%; Pred. No. 0.0036;  
Matches 81; Conservative 45; Mismatches 124; Indels 173; Gaps 17;

33 WHIFSLVILGLFVCGVLMVLYDYNDLSIEDTERENKCVLGPAIVSTGITAVLLV 92  
30 VENFTLVVFGILFAMGVLT-----GNSLVITV 56  
93 LIFVLKRIKLTVELFOITNKALISSAPFLF-----QPLMTFAILL---FFVWLM 139  
57 LARSKRGKRGSTTNLF--ILNLSIDLAVLLFCIPPOATVVALFTVVLGAFICKFIHYFT 115  
140 VAVILSLGTAGAAQV-----MEGQVEYKPLSGIRYMSYHLIGLIWTSEFLA 188  
116 VSMVLSIFTLAMSVDRVVAIVHSRRSSSLRVSRNALGVGFMA----- 160  
189 CQOMTIAAVTCY---FNRSKNDP-----PDHPILSSLSILFFYHOGTVVKGSPFLIS 238  
161 ---LSIAASPAVAYQRLFRHDSNQTFCWEHPNQ-----LHKAYVVCTEVFG 206  
239 VVRIPRIIVMOMALKEQOHGALSRYLFRCCYCCFWCLDKYLLHNLONAVTTAINGTD 298  
207 -----YLLPILLICF-CYAKVNLHLHK----- 228  
299 FCTSAKDAFKILSKNSHFTSINCRGDFIIFLGKVLVVCFTVPGGLMAFNVRAPQVW-- 356  
229 -----LKNMSKSEASKKT-----AQTIVLVVVVFG--ISMLEPHVHIMAE 269  
Db

Tue Jun 1 09:37:58 2004

us-10-063-686-32.rat

Page 9

QY 357 --AVPILLVAFAYVAH-----SFLSY-FETVLDALFLCPAVDLETNDGS 399  
DB 270 FGAFPLTPASFPFRTIHAICLAYSNSVNPITIAFLSENFRAXQVFKRCVNESPHGDA 329  
QY 400 SEK 402  
DB 330 KEK 332

RESULT 14  
US-09-058-333A-5  
Sequence 5, Application US/09058333A  
Patent No. 6368812  
GENERAL INFORMATION:  
APPLICANT: Bard, Jonathan A  
APPLICANT: Borowsky, Beth  
APPLICANT: Smith, Kelli E  
TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/058,333A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52241-E/JPW/XDB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 278 0400  
TELEFAX: 212 391 0525  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-058-333A-5

Query Match 4.7%; Score 108.5; DB 4; Length 346;  
Best Local Similarity 19.1%; Pred. No. 0.0036;  
Matches 81; Conservative 45; Mismatches 124; Indels 173; Gaps 17;

QY 33 WHIFISLVILGLFVCGVLMWLYDYDTNDLSIEDTERENMKCVLGFAIVSTGITAVLLV 92  
DB 30 VENFTITLVVFGILFAMGVL-----GNSLVITV 56  
QY 93 LIFLARRIKITVELFOITNKAISSAPFLF-----OPLWTFAILI---FFWVLW 139  
DB 57 LARSKRGKRSSTTNLF-IINLSIADLAAYLFCIPQATVYALPTVVLGAFICKFIHYFT 115  
QY 140 VAVILSLGTAGAAOV-----MEGQVEYKPLSGIRVMSYHLIGLIWTSFILA 188  
DB 116 VSMVLVSIFTLAAMSVDRYVAIVHSRRSSLRVRNMLLGVPFWA----- 160  
QY 189 COOMTIGAVVTCY-----FNRKNDP-----PDHPILSLSLIFFYHOGTVVKGSPILS 238  
DB 161 ---LSIAMSAPVAYYQRLFHRDSNQTFCWEHWPNO-----LHKRAYVVCFTVFG 206  
QY 229 VRIIRIIVWMONALKEQOGALSRYLFRCCYCCFCWCLDKYLIHLNNAVTTTAINGTD 298

DB 207 -----YLLPLILCF-CYAKVLNHLHKK----- 228  
QY 299 FCTSAKDAFKILSKNSHFTSINCPGDFIIFGKVLVCFVFGGLMAFNRAFOVW-- 356  
DB 229 -----LKNMSKSEASKKKT-----AQTVLVVVVFG--ISMLPHVYIHLMAE 269  
QY 357 --AVPILLVAFAYVAH-----SFLSY-FETVLDALFLCPAVDLETNDGS 399  
DB 270 FGAFPLTPASFPFRTIHAICLAYSNSVNPITIAFLSENFRAXQVFKRCVNESPHGDA 329  
QY 400 SEK 402  
DB 330 KEK 332

RESULT 15  
US-09-595-549-6  
Sequence 6, Application US/09595549  
Patent No. 6511827  
GENERAL INFORMATION:  
APPLICANT: Howard, Andrew D.  
APPLICANT: Cascleri, Margaret A.  
APPLICANT: Smith, Roy G.  
APPLICANT: Sullivan, Kathleen A.  
APPLICANT: Tan, Carina  
APPLICANT: Van der Ploeg, Leonardus H. T.  
TITLE OF INVENTION: GALANIN RECEPTOR GALR3 AND NUCLEOTIDES  
TITLE OF INVENTION: ENCODING SAME  
FILE REFERENCE: 20148PCA  
CURRENT APPLICATION NUMBER: US/09/595,549  
CURRENT FILING DATE: 2000-06-16  
PRIOR FILING DATE: 1998-12-17  
PRIOR APPLICATION NUMBER: US98/26812  
PRIOR FILING DATE: 1997-12-17  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 346  
TYPE: PRT  
ORGANISM: rat  
US-09-595-549-6

Query Match 4.7%; Score 108.5; DB 4; Length 346;  
Best Local Similarity 19.1%; Pred. No. 0.0036;  
Matches 81; Conservative 45; Mismatches 124; Indels 173; Gaps 17;

QY 33 WHIFISLVILGLFVCGVLMWLYDYDTNDLSIEDTERENMKCVLGFAIVSTGITAVLLV 92  
DB 30 VENFTITLVVFGILFAMGVL-----GNSLVITV 56  
QY 93 LIFLARRIKITVELFOITNKAISSAPFLF-----OPLWTFAILI---FFWVLW 139  
DB 57 LARSKRGKRSSTTNLF-IINLSIADLAAYLFCIPQATVYALPTVVLGAFICKFIHYFT 115  
QY 140 VAVILSLGTAGAAOV-----MEGQVEYKPLSGIRVMSYHLIGLIWTSFILA 188  
DB 116 VSMVLVSIFTLAAMSVDRYVAIVHSRRSSLRVRNMLLGVPFWA----- 160  
QY 189 COOMTIGAVVTCY-----FNRKNDP-----PDHPILSLSLIFFYHOGTVVKGSPILS 238  
DB 161 ---LSIAMSAPVAYYQRLFHRDSNQTFCWEHWPNO-----LHKRAYVVCFTVFG 206  
QY 229 VRIIRIIVWMONALKEQOGALSRYLFRCCYCCFCWCLDKYLIHLNNAVTTTAINGTD 298  
DB 207 -----YLLPLILCF-CYAKVLNHLHKK----- 228  
QY 299 FCTSAKDAFKILSKNSHFTSINCPGDFIIFGKVLVCFVFGGLMAFNRAFOVW-- 356  
DB 229 -----LKNMSKSEASKKKT-----AQTVLVVVVFG--ISMLPHVYIHLMAE 269

Tue Jun 1 09:37:58 2004

us-10-063-686-32.rai

Page 10

Qy	357	--AVPILLVAFPAVLVAH-----	SFLSV-FETVLDAFLCPAVDLETNDGS	399
Db	270	FGAFPLTPASFFFRITAHCLAYSNSVNPITIAFLSENFRKAYKQVFKRCVCSBPHGDA		329
Qy	400	SEK 402		
Db	330	KEK 332		

Search completed: May 27, 2004, 16:38:17  
Job time : 24 secs

Tue Jun 1 09:37:59 2004

us-10-063-686-32.rpr

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:32:20 ; Search time 29 seconds  
(without alignments)  
1476.042 Million cell updates/sec

Title: US-10-063-686-32

Perfect score: 2315  
Sequence: 1 MSGRDTTLGLCTIALALSLA.....QDDKSLRNEEGTEQLAVR 445

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	505	21.8	771	2	hypothetical prote
2	257.5	11.1	523	2	hypothetical prote
3	179	7.7	488	2	hypothetical prote
4	177.5	7.7	574	2	hypothetical prote
5	163	7.0	539	2	probable membrane
6	126	5.4	372	2	ubiquinol-cytochro
7	119	5.1	523	2	NADH2 dehydrogenas
8	118.5	5.1	619	2	NADH2 dehydrogenas
9	118	5.1	497	1	NADH2 dehydrogenas
10	117.5	5.1	338	2	NADH2 dehydrogenas
11	115	5.0	868	2	NADH2 dehydrogenas
12	113	4.9	633	2	conserved hypothet
13	111	4.8	493	2	Kup2 Potassium upr
14	110.5	4.8	694	2	NADH2 dehydrogenas
15	110	4.8	495	1	hypothetical prote
16	109	4.7	501	2	NADH2 dehydrogenas
17	108.5	4.7	608	2	ADP-ATP carrier pr
18	108	4.7	474	2	V-type ATPase, sub
19	108	4.7	1783	2	hypothetical prote
20	108	4.7	1917	2	probable voltage-d
21	107.5	4.6	447	2	protein C48A7.1 fi
22	107.5	4.6	627	2	probable O-unit fi
23	107.5	4.6	763	2	hypothetical prote
24	106.5	4.6	501	2	hypothetical prote
25	106.5	4.6	684	2	ADP-ATP carrier pr
26	106	4.6	495	2	NADH2 dehydrogenas
27	106	4.6	497	1	NADH2 dehydrogenas
28	106	4.6	497	1	membrane protein L
29	106	4.6	677	2	hypothetical prote
					sulate transporte

30	105.5	4.6	332	2	B82876	oligopeptide trans
31	105	4.5	782	2	T25925	hypothetical prote
32	104.5	4.5	342	2	T29245	hypothetical prote
33	104.5	4.5	475	2	E37346	MG294 homolog A05-
34	104.5	4.5	492	2	E58931	NADH2 dehydrogenas
35	104	4.5	245	2	B60944	ubiquinol-cytochro
36	104	4.5	445	2	E22845	hypothetical prote
37	104	4.5	544	2	C96943	ubiquinol-cytochro
38	103.5	4.5	250	2	C60944	characterized me
39	103	4.4	446	2	S01187	NADH2 dehydrogenas
40	103	4.4	461	2	S60253	sel-12 protein - C
41	103	4.4	518	2	A53207	probable folate tr
42	103	4.4	602	2	D90557	hypothetical prote
43	103	4.4	1128	2	T04922	hypothetical prote
44	102.5	4.4	306	2	A33960	cytochrome ca3 ox
45	102	4.4	383	2	S55594	G protein-coupled

#### ALIGNMENTS

##### RESULT 1

T16254  
hypothetical protein F35C8.7 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C/Accession: T16254  
R.Wu, X.

submitted to the EMBL Data Library, November 1995  
A/Description: The sequence of C. elegans cosmid F35C8.

A/Reference number: Z18486  
A/Accession: T16254

A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA

A/Residues: 1-771 <MW>  
A/Cross-references: EMBL:U40941; NID:G1072184; PID:G1072187; PIDN:AAA81709.1; CESP:F35

A/Genetics:  
A/Genes: CESP:F35C8.7

A/Introns: 87/2; 143/2; 238/1; 289/2; 428/2; 629/3; 678/1; 712/1; 736/1

Query Match 21.8%; Score 505; DB 2; Length 771;  
Best Local Similarity 28.9%; Pred. No. 2.9e-33;

Matches 131; Conservative 95; Mismatches 180; Indels 48; Gaps 13;

QY	10	LCIALALSLAMTTPRTITLVAHITISLVITGLVVCGLVW-WLYDYDT--NDLSI	64
DB	318	LIFPAAGILSFTWVIMRLGSLIILVLVALGAGSWMKWTLLKTGAIDYVSF	377
QY	65	E--LTERENMKCVLGFIVSTGTAVALVLFVLRIRKLTVELPQITNKASAPFL	122
DB	378	HPADAYEMFTLVVAIAISVLLFLVILFIRIRISACILSESSKAISSMMSTL	437
QY	123	FOPLMTFAILIFVVLVAVLISGTNG--AAQVME-GGQVEYK-----PLSGIRY	170
DB	438	LFPLFPLFHLHGVFALMGSIAIMWLASGQEVCRILKENTGVV-YNTSTKCDCTAKVTCQV	496
QY	171	M-----W--SYHLGLIWTSEFLACQQTINGAVVTCFENSK-NDPPDHILS	217
DB	497	VGIEKESSTIYLVQVTLNFAFWISCVTLGDLALGARASVYAWDKHDVFTFVIR	556
QY	218	SLSLIFFYHOGTVKSGFLISVNRIPRIIVTMQNALKEOHGALSRYLFRCCYCCFWCL	277
DB	557	ALNRAIRNKGISAFGSLIIAIVKIIKIVLEVIDHKLGRSQNKAVKFM-CLKCCWCL	615
QY	278	DKVLLHNNQAYTTAANGDPCISADAKILSKSSHSTISNCFDPIIFLGLVVC	337
DB	616	EVFFKFLKNAVIMTALVYGNPFSSADSFLITRNIVRTVVAKVAGILLFLGKSLI--	673
QY	338	FTVEGGLMAFNVRAPV-----WAVPLLVAFPAVVAHSEFLSVFETVDALF	386
DB	674	-TLGAGILSFYFSGRNVVGVPRVDLYYFVPIVIVVISYMMADLFFVIVYEMADYTF	732
QY	387	LCFAVVDLETFDGSSEKRYFMDQEFSLTVKASNTL	420

Db 733 ICFLEDSQNDGSLERFFWSEKLFILGNKDI 766

## RESULT 2

T05692

hypothetical protein F20M13.200 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999

C/Accession: T05692

R/Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.H.

submitted to the Protein Sequence Database, February 1999

A/Reference number: 215420

A/Accession: T05692

A/Molecule type: DNA

A/Residues: 1-523 <BEV>

A/Cross-references: EMBL:AL035540

A/Experimental source: cultivar Columbia; BAC clone F20M13

C/Genetics:

A/Map position: 4

A/Introns: 264/2; 339/2; 391/1; 437/3

A/Note: F20M13.200

Query Match 11.1%; Score 257.5; DB 2; Length 523;

Best Local Similarity 24.6%; Pred. No. 2.7e-13;

Matches 100; Conservative 75; Mismatches 160; Indels 71; Gaps 16;

QY 6 TILGCIATLALSLAMFFFRITTLVHIFISLVILGLFVCGVLMWLYDYNDISIELDTERENKCVLGA 65  
 Db 145 TLVWVLISVPCFESVLLKRYTKQIYACPLVLEPIF-NTYMWACTLSGSCDA 203  
 QY 66 LDTRENNKCVLGAIVSTGTAVALLVLPVKRKIKLTVELFOITNKAISSAPPLFQF 125  
 Db 204 LPLAVRIIVLVFVFPII-----GIIVITIVNWHIDITLQISVSDALS-----KN 251  
 QY 126 LMTFAILIFFWLWAVLISLGTAGAQVME-GGQ--VEYKPLSGIRFMSYHLGLMT 182  
 Db 252 LKLFVVL-----PLITLGFNKKFVPREIDQYFCWEKEDSVWPAYVALILNIMS 302  
 QY 183 SEFIIACQOMTIGAVTVCYFNRSKNDP-PDHPILSSLSILFFYHQGTGVKGSFLISVVR 241  
 Db 303 LAVWEMQVYVIGALIAQWYF--SKEDSIPKKCIRSRNRFQSGFTIGVSGLLICIVR 360  
 QY 242 IPRITVMTQNALKEQOGALSRVYFRCCYCCFGLDKYLHMONATTTAINGTDFCT 301  
 Db 361 VRAIV--DNARENTQIVWV-LKCC-----ANALLGEAYCT 396  
 QY 302 SAKDAFKILSKN--SSHF-----TSINCFGDFIIFLGVLVCFVFGGLMAF--NYNR 351  
 Db 397 SAKTYELRRRLISAVVEVETSTRI-LTGIVFLSAAYAVATWAVLKGVSNLGIDSVYV 455  
 QY 352 AFQWAVPLLVAFAYVAHSFSLSVETVLDALFLCAVLENTND 397  
 Db 456 AVLAVWLVLLVILAFVH-----VLDVDVIDTVVCAIDRDKGD 493

## RESULT 3

C66385

hypothetical protein F20J.7 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C/Accession: C66385

R/Thelloglis, A.; Becker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

Nature 408, 815-820, 2000

A/Authors: Hunter, J.U.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, C.J.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: 266141 MIM 21016719, PMID 11300712

A/Accession: C66385  
 A/Status: Preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-488 <STO>  
 A/Cross-references: GB:AE005172; NID:G10092332; PIDN:AA12743.1; GSPDB:GN00141  
 C/Genetics:  
 A/Map position: 1

Query Match 7.7%; Score 179; DB 2; Length 488;

Best Local Similarity 21.1%; Pred. No. 6e-07;

Matches 85; Conservative 62; Mismatches 146; Indels 110; Gaps 18;

QY 21 MFFPRITTLVHIFISLVILGLFVCGVLMWLYDYNDISIELDTERENKCVLGA 80  
 Db 128 MKSVHILTYLA-----VSVLCFWCCLFF-----GGA 157  
 QY 81 IVSTGTAVALLV-----LVLRKRIKLTVELFOITNKAISSAPPLFQPLMTFAI 131  
 Db 158 FAVGSLQFLYVISVDRPLFTWLVRKALKVWGPKYVVA-----HAFIV 205  
 QY 132 LIFPVV-LWVAVLISLGTAGAQVMEGQVEKPLSGIRFMSYHLGLMTSEFLAQ 190  
 Db 206 VMLKMSLV-----STGAAGVVAASMGDE-----GRWLVVLSLFWTGAVALCNTV 253  
 QY 191 QMTIAGAVTVCYFNRSKNDP--PDHPILSSLSILFFYHQGTGVKGSFLISVVRIPRIIV 247  
 Db 254 HVIIVGAVFVHLFHCQESSSLPESLVDLSRYAVTTSFGISVYGSFLTAIR----- 307  
 QY 248 MNQNLKQOGHALSRVYFRCCYCCFGLDKYLH-----NQAATTATNGTDF 299  
 Db 308 ----TLRWEIRFRSKICGNBCLC--CVD-FLFELVETLVRFKVAVVOIAVYKGF 359  
 QY 300 CTSKADAFKIL-SKNSHFTSINCFGDFIIFLGVLVCFVFGGLMAFNARQVMA- 357  
 Db 360 NKSADAMLEFQSTGEALVAADCSS-AVLNG-----TIFGLIT--GSCIGIMAM 408  
 QY 358 ----VPLLVAFAYVAHSFSL--VEETVLDALFLCAVND 392  
 Db 409 IKYSDRVIVASTAMLMGVLVGLQWVESAIVTSIYICFAD 451

## RESULT 4

T41068

hypothetical protein SPCC1682.11c - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 04-Mar-2000

C/Accession: T41068

R/Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A/Reference number: 221968

A/Accession: T41068

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-574 <MUR>

A/Cross-references: EMBL:AL031525; PIDN:CAA2067.1; GSPDB:GN00068; SPDB:SPCC1682.11c

C/Genetics:

A/Experimental source: strain 972h-; cosmid c1682

A/Gene: SPDB:SPCC1682.11c

A/Map position: 3

A/Introns: 5/1; 358/2; 538/2

C/Superfamily: Schizosaccharomyces pombe hypothetical protein SPCC1682.11c

Query Match 7.7%; Score 177.5; DB 2; Length 574;

Best Local Similarity 21.4%; Pred. No. 9.4e-07;

Matches 91; Conservative 76; Mismatches 159; Indels 99; Gaps 20;

QY 12 IIALAIALSLAMFFFRITTLVHIFISLVILGLFVCGVLMWLYDYNDISIELDTERE 71  
 Db 192 IICMMISVIMLFCVLAIPRFVLFYLLASVPLTWFAFAYVL-----KASRIHLETSIQ 243  
 QY 72 NKKCVLGAIVSTG---TAVLVLLFVLRKRI--KLVEYFOITNKAISSAP-----FL 121  
 Db 344 NVNVI---LWVAVLISLGTAGAQVMEGQVEKPLSGIRFMSYHLGLMTSEFLAQ 190

QY 122 LFOPLMTFALLIFFWLWVAVLISLGT--AGAAQVMEGGQVEYKPLSGIRYK--SYHL 176  
 Db 298 SF--LRSFYLLIFWVLFRLRFLRSTLWGSVWLP-----RSSWVLASFYS 343  
 QY 177 IGLWTFSEFLAACQMTIAGAVTCTCFNRKNDPPDHPILSLIFPNH-----OGTVV 231  
 Db 344 LHLMLCTFFHAIQCALISSIVSQWFFYR---DTKSATKTNLVSHFFTHVVSNOYGLCA 400  
 QY 232 KGSFLLISVARIPIIIVMTQNALKEQGHALSRYLFCYCCGCEWCLDKYLLHNLQNAVYT 291  
 Db 401 FSSFLVITVTP--LHFLPTWL-----RHVGRIVY-----YMFSTKTSASYT 440  
 QY 292 T-----AINGDFCTSAKDAFKI-----LSKSSHFTSINCGRDEIIFLGKTLVVC 337  
 Db 441 SPLTLAVASISYSPYMAASAKALYQIEQNRVGLRRSYFBSK-----ITLLAARSL-- 492  
 QY 338 FTVFGGLMAFNY---NRAQVWAVPLLVAFPAVLVASHFLSVFETVLDALFLCPAVDL 393  
 Db 493 -AIGGVTSWNVYSIHENGVEYGYIVG-LIGGFILWLLIGABGLSMIVDALLIGSIDDI 550  
 QY 394 ETNDG 398  
 Db 551 SSCQG 555

## RESULT 5

S67049  
 probable membrane protein YOR161C - yeast (Saccharomyces cerevisiae)  
 N/Alternate names: hypothetical protein O3568  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002  
 C/Accession: S67049  
 R/Bordome, R.; Camassee, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarasov, I.A.; Wlasek  
 submitted to the Protein Sequence Database, July 1996  
 A/Reference number: S67032  
 A/Accession: S67049  
 A/Molecule type: DNA  
 A/Residues: 1-539 <BOR>  
 A/Cross-references: EMBL:Z75069; NID:g1420398; PID:e252377; PID:g1420399; GSPDS:GN00015;  
 A/Experimental source: strain S288C  
 C/Genetics:  
 A/Gene: MIPS:YOR161C  
 A/Cross-references: SGD:S0005687  
 A/Map position: 15R  
 C/Keywords: transmembrane protein  
 F/84-100/Domain: transmembrane #status predicted <TM1>  
 F/159-175/Domain: transmembrane #status predicted <TM2>  
 F/183-199/Domain: transmembrane #status predicted <TM3>  
 F/235-251/Domain: transmembrane #status predicted <TM5>  
 F/278-294/Domain: transmembrane #status predicted <TM6>  
 F/337-353/Domain: transmembrane #status predicted <TM7>  
 F/372-388/Domain: transmembrane #status predicted <TM8>  
 F/437-453/Domain: transmembrane #status predicted <TM9>  
 F/474-490/Domain: transmembrane #status predicted <TM10>

Query Match 7.0%; Score 163; DB 2; Length 539;  
 Best Local Similarity 21.1%; Pred. No. 1.3e-05;

Matches 94; Conservative 73; Mismatches 160; Indels 118; Gaps 20;

QY 10 LCILA-----LALSLAMPTFRFITLLVHFISLVIIIGLIFVCGVLMWLVYDYNLDI 64  
 Db 136 VCIILVSVGLTRICRIFPKPI-----YCGAV----- 164  
 QY 65 ELDTRENNKCVLGAIV-----STGITAVLVLIV-----LRKRIKLTVELFOI 110  
 Db 165 -----INLVASLGRIMVMSLRYSAGI--VFVFTMTACWYGMKRSRIPLSVAVALKV 216  
 QY 111 TKKAISSAPFLIFQGLMTFAILIFFWLWVAVLISL-----GRAGAAQVMEGGQVEYK 163  
 Db 217 VVDANKKCPQIFTVFVGLAVASAGFLFSAVYATVYIKYIDPNSSNGGCDV--SGGSCSHS 275

QY 164 PLSGIR--YMSYHLIGLWTFSEFLAACQMTIAGAVTCTCFNRKNDPPDHPILSL 219  
 Db 276 KLIGLVVVPFCGYI-----SEVIRNVJHCIVSIGVFGSWYMSKSDQGMPPARFAGL 329  
 QY 220 SILPFYHQGVKSPILSVARIPRIIIVMTQNALKEQGHALSRYLFCYCCGCEWCLD 278  
 Db 330 KRAMTYSGSCIFCSGLVALIDLRLQILQMTNRHVYTSGGQILA---IQILFWFDWITG 386  
 QY 279 --KYL--HLNQNAVYTTAINGTDFCTSAKDAFKILSKN-----SSHFTSI-----NCFG 324  
 Db 387 FLKWLAEYFNHYAASFALYKPYLRAKETWYLRKGMALINDMINIALGLFSMFA 446  
 QY 325 DEIFILGKTVLVCT-----VFGIWAFFNNRAFOVW-----AVPLLVAFAYL- 369  
 Db 447 SYMTALFTFLYLRFTSPQVNSNGAYNGALMAFSVALQICNATEAIRSGTATFVALG 506  
 QY 370 -----VAHSEFLSVFETVLDAL 385  
 Db 507 NDPEVFNHSHYPRDDEIFRAVPDL 531

## RESULT 6

S52054  
 ubiquinol-cytochrome c reductase (EC 1.10.2.2) cytochrome b - Trypanoplasma borreli mi  
 C/Species: mitochondrion Trypanoplasma borreli  
 C/Date: 14-Jul-1995 #sequence\_revision 01-Nov-1996 #text\_change 03-Jun-2002  
 C/Accession: S52054  
 R/Lukes, J.; Arts, G.U.; van den Burg, J.; de Haan, A.; Opperdoes, F.; Sloof, P.; Benn  
 EMO U. 13, 5086-5098, 1994  
 A/Title: Novel pattern of editing regions in mitochondrial transcripts of the cryptobi  
 A/Reference number: S52053; PMID:95045401; PMID:7525275  
 C/Accession: S52054  
 A/Molecule type: mRNA  
 A/Residues: 1-372 <LUK>  
 A/Cross-references: EMBL:U11684; NID:g565141; PIDN:AAA73455.1; PID:g565142  
 A/Genetics:  
 A/Gene: cytb  
 A/Genome: mitochondrion  
 A/Genetic code: SGC6  
 C/Superfamily: cytochrome b, cytochrome b homology; cytochrome b6 homology; plastoquin  
 C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrio  
 F/10-343/Domain: cytochrome b homology <CBH>  
 F/110-211/Domain: cytochrome b6 homology <CB6>  
 F/1225-343/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>  
 F/183,184/Binding site: heme iron (His) (axial ligands) (low potential) #status predict  
 F/197,198/Binding site: heme iron (His) (axial ligands) (high potential) #status predict

Query Match 5.4%; Score 126; DB 2; Length 372;  
 Best Local Similarity 20.7%; Pred. No. 0.0094;

Matches 92; Conservative 64; Mismatches 124; Indels 164; Gaps 24;

QY 10 LCILA-----LALSLAMPTFRFITLLVHFISLVIIIGLIFVCGVLMWLVYDYNLDI 61  
 Db 17 LCILLISGDLFRYVGLGFGNIGWAIQI-----LVGIC--LSMFFRCIIPQWMTFTLL 68  
 QY 62 LSIELDTRENNKCVLGAIVSTGI--TAVLVLIVLRRRIKLTVELFOITKAISSAP 119  
 Db 69 IHLFFD-----IGFIIRSLHIIFTSILVFLVLIYHIHKVIFLCIIFDS----- 111  
 QY 120 FLFQPLMTFAILIFFWLWVAVLISLGT--TAGAAQVMEGGQVEYK-----PLSGIRYK 172  
 Db 112 ---MLVMEFGFLIFILIAI--IGYLPCTSNYSWGLTVFSNLTATPILGI----- 160  
 QY 173 SYHLIGLWTFSEI-----LACQMTIAGA--VTCYFNRSKNDPPDHPIL 216  
 Db 161 --YICQWICSEFINFTLLKLSIHIFLPVLLFLIGAHFVLHYF-----L 206  
 QY 217 SLSLSI-----PFYHQGVKSPILSVARIPRIIIVMTQNALKEQGHALSRYL 267  
 Db 207 SSDGLDRPFYERFFFLYLRDLPLI--INILCFLLYYI----- 247  
 QY 268 RCYCCFWCLDKYLLHNLQNAVYTTAINGT-----FCTSAKDAFKILSKNS 314



Db 248 ----CIW-----FEVHEBSWIVDTLTKSDKILPEWFFLSFGFLKSPDKF----- 292  
 Qy 315 SHFTSINCFDGFIFLGLKVLVVCFTVF-----GGMAFNYNRAFOVMAVPLLVAFAVLY 370  
 Db 293 -----MGFLPLF-----VLCFALPEFLINCLILITICRSSLIM-MSLSILVLFY-YLC 337  
 Qy 371 AHSFLSVFETVLDALFLCFPAVDLE 394  
 Db 338 VGGFLSLY-----VVLCFPLME 355

## RESULT 7

T11916  
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Prototheca wickerhamii mitochondrion  
 C:Species: mitochondrion Prototheca wickerhamii  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 03-Jun-2002  
 C:Accession: T11916  
 R:Wolff, G.; Plante, I.; Lang, B.F.; Kueck, U.; Burger, G.  
 J. Mol. Biol. 237, 75-86, 1994  
 A:Title: Complete sequence of the mitochondrial DNA of the chlorophyte alga Prototheca wickerhamii  
 A:Reference number: 217373; MUID:94180393; PMID:8133522  
 A:Accession: T11916  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-523 <MOL>  
 A:Cross-references: EMBL:U02970; NID:9467843; PID:9467848; PIDN:AA012635.1  
 A:Experimental source: strain HB-8  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Note: nad4  
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4  
 C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 5.1%; Score 119; DB 2; Length 523;

Best Local Similarity 18.1%; Pred. No. 0.049;  
 Matches 76; Conservative 63; Mismatches 163; Indels 118; Gaps 15;

Qy 11 CIALALSLAMWTFRRFTTLVAFITSLVIGLFFVGVLMWLYDYNDLSIEDTER 70  
 Db 151 CIALVLETMLTVFSVLDLLEFFIFESVLIPIFIIGV-WG-----SRER 196  
 Qy 71 ENMKCVGLPAIVSTGITAVALVILFVLRKIKLTVELFQTNKAISSAPLLFOPLMTPA 130  
 Db 197 KIRAYOFFLTVLFGSVMLAILIIFQGTLDIEMLYLSDSETQCILMLAFASFA 256  
 Qy 131 IL-----IFW-----VLMVAVLSLGTAG---AAQWEGGVEXKPSGI 168  
 Db 257 VKVMVPMVHVLPEAHVBAFTAGSVIILKIKGTGFLRFSPLPYACIYTPP--- 313  
 Qy 169 RYMSYHLIGLWTS-----EFLACQOMTIAGAVTCYFNRSKNDPPDPIILSL 219  
 Db 314 --IYMSVIAIVYTSCTIRQIDKLIIVSSVAMNFVILGLFSQNT----- 359  
 Qy 220 SILFFYOGIVVKSFLISVVRIRIIVMWQNLKQOQAGLSRYFRCCYCCPWLCK 279  
 Db 360 -----QG--TESILMT-----SHGLVSPALFLCGVLY----- 387  
 Qy 280 YLHLNQAAYTTAINGDPTCSAKDAFKILSKNSHRTSINCP-----GDPIIFLG 331  
 Db 388 -----DRHKTRLRLRYSGCGQMPFALL--FVEFTMANISLPTGSPGGEFLVFIG 437  
 Qy 332 KVLVVCPLVFGGLAFAFNRAFOVMAVPLLVAFAVLYAH-SLSVFEVTLDALFLCEA 390  
 Db 438 SYQNNSFVAFCAATGMVIGAAVYALMLCNRLI---YGVSKPDPIINTWSDVNRREFMFA 492

## RESULT 8

T11314  
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - Pedinomonas minor mitochondrion  
 C:Species: mitochondrion Pedinomonas minor  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 03-Jun-2002  
 C:Accession: T11314  
 R:Turne, M.; Lemieux, C.; Burger, G.; Lang, B.F.; Otis, C.; Plante, I.; Gray, M.W.

submitted to the EMBL Data Library, December 1998  
 A:Description: The complete mitochondrial DNA sequences of Nephroselmis olivacea and F

A:Reference number: 217261  
 A:Accession: T11314

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-619 <TUR>

A:Cross-references: EMBL:AF116775; NID:94378766; PID:94378768; PIDN:AA019666.1

C:Genetics:  
 A:Genome: mitochondrion

A:Note: nad5  
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 5.1%; Score 118.5; DB 2; Length 619;  
 Best Local Similarity 22.0%; Pred. No. 0.064;  
 Matches 105; Conservative 73; Mismatches 160; Indels 139; Gaps 24;

Qy 7 ILGLCILALSLAMWTFRRFTTLVAFITSLVIGL-----LVTLG 43  
 Db 124 ILVAVSEVALTLFLGWDGIGITSYLLINFWYSKIQSKSAIKAVFLNRIDSPFLMLG 183  
 Qy 44 LFFVGVLMWLYDYNDL-----SIEDTERENMK-CVLGPAIVSTGITAVALVILFVLR 98  
 Db 184 LTF-----YLFSGDDLFLISLSVFEKQIVKYLIAVLIAISAKAQILHVM-LP 234  
 Qy 99 KRIKLTVELFQTNKA--ISSAPLLFOPLMTPALIFP---WLMVAVLSL--GTAG 150  
 Db 235 DAIEAPTPVSSLHATATVAGVYLIR--LSFLDLPDSNFIIVIGIITSPLAGIG 292  
 Qy 151 AAQWEGGVEXKPSGISRYMSYHLIGI-----WTSEFILAQQOMTIAGAVTCY 202  
 Db 293 FMOPTKRIAYSTCSQIGM--FYALGISLSDPSYLFHVHGPKCMFLAAGFIITI 350  
 Qy 203 FNRSKNDPPDHP-----ILSSSLT--FFYHGTGVKSGFLISVVR--PRIIVM 248  
 Db 351 FN--EQDIRKYSGLNPLNSFLIAFVLSNLSLGIFFAGFYSKELLISGI----- 400  
 Qy 249 WQNLKQOQAGLSRYFRCCYCCPWLCKYLHLNQAAYTTAINGDPTCSAKDAFK 308  
 Db 401 YTN-----FWSL-----LSVLAFTTCLVY-----IKSILL 427  
 Qy 309 IL--SKNSHRTSINCFGD--FIIFLGLVAVVCFVFG-----GLMAFNYNRAFOVMA 357  
 Db 428 VLSGSPNNMSPFSITTYTNLIVILISLVVILNLFPGPMITEQIKLMDIYVNS----- 482  
 Qy 358 VELLVAFPAYVANSFLSVFETVLDALFLCPAVDLEINDDSSSEKPYMDDEFLSV 414  
 Db 483 -----VPS-----IKSNFLFEHYVYVNIILFVLSVFLISLWNSYIFPFSKLSFI 530

## RESULT 9

S53834  
 NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Acanthamoeba castellanii mitoc  
 C:Species: mitochondrion Acanthamoeba castellanii  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
 C:Accession: S53834  
 R:Burger, G.; Plante, I.; Loneragan, K.M.; Gray, M.W.  
 J. Mol. Biol. 245, 522-537, 1995  
 A:Title: The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba castellanii: cc  
 A:Reference number: S53825; MUID:95147275; PMID:7844823  
 A:Accession: S53834  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-497 <BUR>  
 A:Cross-references: GB:U12386; NID:9562028; PIDN:AA011826.1; PID:9562038  
 A:Experimental source: strain Neff; ATCC 30010  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
 C:Genetics:  
 A:Genome: mitochondrion  
 A:Note: nad6  
 C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4  
 C:Accession: SGC6

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 5.1%; Score 118; DB 1; Length 497;  
Best Local Similarity 19.5%; Pred. No. 0.056;  
Matches 90; Conservative 65; Mismatches 143; Indels 164; Gaps 22;

```

QY 11 CIALALSLAMMTFRITTLVHIFSLVILG-----LIFVCGV 50
DB 140 CVALD-----VEFYIFESILIPMFILIGWGRORHAYVOLFPPYLLGSLMLLGI 193
QY 51 LMMVYDYTDLSIELDERENKCVLGAIVSTGIALVLLVIFLRRKIKLVELPQI 110
DB 194 LVYISHQTDIVLVNTN-----FSFYRLILMASPF-- 226
QY 111 TNKAISAPPLFQPLMTFAILLFW-----VLWVAVLLSGTAGAAQVMBG 157
DB 227 -----FACVCKVLPFPFHI-----MLPEAVPEAPTVGSVILAGVLLKLGTYGL----- 269
QY 158 GQVEYKPLSGIRYV-----MSYHLGLTWSEFILAQQOMTAGAVTCYFRSKNDP 211
DB 270 -----LRFVIFPCDAITFFLPVYT-----LCMLGITTCSTIRQVDLK 310
QY 212 DHDILSLILFFYHOGTVKGSFLISVVRIPRIIVMVMONALKEQOHALSRYLEPCY 271
DB 311 KVIAYASVS-----HMSFVILGLFTSNIGIGSVFLML-----SHGIVSGLPFCIG 358
QY 272 CCWCMDKY---LHINQNAVTTAINGDPCISADAFAKILSKNSHFTSINCQDPFI 328
DB 359 CVY---DRYKTRILRYSGLVSTMP--FSLC-----LTLILSNISFPGTSFGEFVI 408
QY 329 FLG-----KVLVCFVFGGMAFNVRAPQVAVPPLVAVFAVLAHSPF-----S 376
DB 409 LGLFENNHPALIAIFSTI-----LNAVSIWLYNRKI---FRLVAVNYILRPSDS 458
QY 377 VERTVDALFLCPAVDLETNDG-----SEKPYMDQEFLLS 413
DB 455 KKEFVVGFI-CPITTLFGKGSYIILIEAPLYV---YLSF 496

```

# RESULT 10

550339  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Katharina tunicata mitochondrion

C:Species: mitochondrion Katharina tunicata

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 03-Jun-2002

C:Accession: S50339

R:Boore, J.L.; Brown, W.M.

Genetics 138, 423-443, 1994

A>Title: Complete DNA sequence of the mitochondrial genome of the black chiton, Katharina

A:Reference number: S50327; MUID:95129806; PMID:7828825

A:Accession: S50339

A:Molecule type: DNA

A:Residues: 1-338 <BOO>

A:Cross-references: EMBL:U09810; NID:9557273; PIND:AA48374.1; PID:9335326

A>Note: the authors translated the initiation codon GTG for residue 1 as Val

C:Genetics:

A:Gene: ND2

A:Genome: mitochondrion

A:Genetic code: SGC4

A:Start codon: GTG

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 2

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 5.1%; Score 117.5; DB 2; Length 338;  
Best Local Similarity 21.6%; Pred. No. 0.042;  
Matches 85; Conservative 52; Mismatches 120; Indels 137; Gaps 20;

```

QY 22 MPTFRITTLVHIFSLVILGLFVCGVLMVLYDYNDLSIELDERENKCVLGPAL 81
DB 1 MNPFP-----VGLIFILFFGTLPSLSIHW-----FGVWGLELN-----LMGF-- 41
QY 82 VSTGTAVALVILFVLRKRIKLT---VELFOITNKAISAPPL--LFPPLWTFAL----- 131
DB 42 -----IPVWVQKSTSEETESGVKYLIV--QAVGSAFLFLGLMLMNNPFCWELNF 89

```

```

QY 132 -----LIFPWLWVAVLLSGTAGAAQVMEGQVEYKPLSGIRYVMSYHLI-GLIWT 182
DB 90 FSGFSGXSLIIF-----GLMKLGA-----P-----HWVSVVAGLSM 126
QY 183 SEFIIACQOMTAGAVVTCYFRSKNDPDPHILSLISILFFYHOGTVKGSFLISVRI 242
DB 127 SNFLLITVQ-KIAPLFVWCYIN-----LSSFLLITLIVLSSIFGVGVGVMQTSV 175
QY 243 PRII-----VMMONALKEQOHALSRYLEFRCCYCCFMDLVYLLHINQNAVTTAINGD 298
DB 176 PALIYSSILHMGWMLKXASAGMSIFPFYFPGFIIGFAVLMGLDE----- 223
QY 239 FCTSAKAFKILSKNSHFTSI---NCFG-DPIIFLGKVLVCFVFGGLMAFNVRAPQ 354
DB 224 -----SFNMSCFSVYVWNSYSRFLVFM-----LLSIGL----- 254
QY 355 VNAVPLLVAFPAVLVAHSPFISVEFYLDAFLC 388
DB 255 ---PPLDGFQKMLVAVSLSLGNLVLSITLVC 284

```

# RESULT 11

B89897  
conserved hypothetical protein SA1085 (imported) - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: B89897

R:Kurdoz, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Og

ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.

C:Shiba, T.; Hattori, M.; Ogawara, N.; Hagiashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: B89897

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-868 <KUR>

A:Cross-references: GB:BA000018; PID:913701042; PIND:BA842337.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA1085

Query Match 5.0%; Score 115; DB 2; Length 868;  
Best Local Similarity 18.6%; Pred. No. 0.17;  
Matches 73; Conservative 63; Mismatches 140; Indels 116; Gaps 16;

```

QY 8 LGLCIALALSLAMMTFRITTLVHIFSLVILGLFVCGVLMVLYDYNDLSIELD 67
DB 197 IGRIVAILTLPSNFYS-----YQALIIIG---CYLVKRIIFYKXDIV----- 239
QY 68 TERENKCVLGAIVS-----TGITVL-----LVILFVLRKRIKLTVEL 107
DB 240 SRQKFLCVASATVLSVSVFGLFTGISAFLENDKQNNVNDIPILTPDHYHFFESDG 299
QY 108 FOITNKAISAPPLFQPLMTFAILLFWVAVVLLSGTAGAAQVMEGQVEYKPLSG 167
DB 300 FYIITSLITVALLSPGLVYFRYRLPAITWTLFPGSL-----SQFDSA---FNGFSF 351
QY 168 IRYWASHYHLGLTWSEFIIACQOMTAG-----AVTGYFRSKNDP----- 210
DB 352 PERRWVY-----LIALSSALCGFIQHLSTLNMKYILRTIPVSIALLVYLL 400
QY 211 -PDHPILSSILFFYHOGTVKGS-----FLISVVRIPRIIV-----Y 249
DB 401 SPFHPLALYGIILWVAVILKFSIMRYKGLTVALVILVIMQCVLLDNNKMAIKP 460
QY 250 MONALKEQOHALSRYLEFRCCYCCFWCLDKYLHINQNAVTTAINGDPCISAKDAFKI 309
DB 461 QQSISTLKQHDYHSN-----VNLQIKINQNA--TGFPRIDYMSD---YAL 503
QY 310 LSKNSHFTSINCF-----GDPIIFLGKVLV 336

```

Db 504 NSPFIYHNGISLYSIFNGLIKYDKTLQI 535

# RESULT 12

Kup2 Potassium uptake protein [imported] - Sinorhizobium meliloti (strain 1021) magaplas  
C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C/Accession: G95385

R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows  
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
A/Reference number: A95262; MUID:21396509; PMID:11481432

A/Accession: G95385

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-633 <KIR>

A/Cross-References: GB:AE006469; PIDN:AAK65649.1; PID:G14524136; GSPDB:GN00165

A/Experimental source: strain 1021, megaplaemid psyma

R/Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaulc, P.; Vandenberg, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A/Title: The complete genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A:Genome: kup2

A:Genome: plasmid

Query Match

Best Local Similarity 4.9%; Score 113; DB 2; Length 633;

Matches 49; Conservative 39; Mismatches 67; Indels 68; Gaps 9;

Db 3 GRDTIIGLCTIALSLAMMFFRTITTL-----VHIFSLVIG-----LHVV 47

Db 54 GRBEVIGLVSLLV-WITTAIVTITKYVLFILRADNDGEGTLSTLALLKKTKYPTLMFF 112

Qy 48 CGVLMWLY-----DYTNDSIELDERENKCVLGFAIVSTG-----ITAVLVLFV 96

Db 113 AGVLGALFLIGDAMITPALS-----LSNVEGLKLVAPALADVVLISVIIILLRA 164

Qy 97 LRRIKLTVELFQITNKAISSAPFLFQPLMTFALIFFWLVAVVLSLGTGAAGVME 156

Db 165 VQGR-----GTGAVSVFFQPIITLWPFLLMAAAGVNHID 198

Qy 157 GGVV--EYKPLSGIRYMSYHLGLTWSEFLLACQMTIAGA 197

Db 199 DLALISAFNPLNAGIFLNMAGLIGFT-----VLGALFLVTGA 236

RESULT 13

S78183

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Reclinomonas americana (ATCC 503

C/Species: mitochondrion Reclinomonas americana

A/Variety: ATCC 50394

C/Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 03-Jun-2002

C/Accession: S78183

R/Lang, B.F.; Burger, G.; O'Kelly, C.J.; Cedergren, R.; Golding, G.B.; Lemieux, C.; Sank

Nature 387, 493-497, 1997

A/Title: An ancestral mitochondrial DNA resembling a eubacterial genome in miniature.

A/Reference number: S78127; MUID:97311393; PMID:9168110

A/Accession: S78183

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-493 <LAN>

A/Cross-References: EMBL:AF007261; NID:G2258325; PIDN:ADD11916.1; PID:G2258382

A/Experimental source: ATCC 50394

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1997

C/Genetics:

Query Match

Best Local Similarity 4.8%; Score 110.5; DB 2; Length 694;

Matches 88; Conservative 73; Mismatches 164; Indels 105; Gaps 20;

A:Gene: nad4

A:Genome: mitochondrion

C/Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation

Query Match

Best Local Similarity 4.8%; Score 111; DB 2; Length 493;

Matches 85; Conservative 53; Mismatches 128; Indels 124; Gaps 21;

Db 10 LCTIALSLAMMFFRTITLLVHIFSLVILGLFVCGVLMWLYDYTNDSIELDTE 69

Db 120 ICEFLDLALLIMIFC--VLDELVFYIFPESVLIMFVIGV-WG-----SRE 163

Qy 70 RENNKCIVGPAIVSGITAVLVILFV-----RRIKLTVELFQITNKAISSAPFL 122

Db 164 RYVRAVWLFVYTFPG-SLMMLALWVYFDAGTTDQVLTFFSQRQLWLAFFIS 222

Qy 123 PQ--PLWTFALLIFW-----VLWVAVLLSCTAG--AAQWEGQVEY 162

Db 223 FAIKIPWVFFHV---MLPEAHVAPRAGSVLGLVLLKGGYGLRPSIMPEATVYF 278

Qy 163 KPLSGIRWMSYHLGLTWSEFLLACQMTIAGAVVTCYNRKNDPDPHPIISSIL 222

Db 279 TPL-----VYMSITAIYTS--LTLRQIDLKRIIA-----YSSVA-- 313

Qy 223 FFYHGVVAKGSFLISVVRIPRIITWYQNALKEQHGALSRYLFRCCYCCFCLDKY-- 280

Db 314 --HNFTIGMFLMNOGLESLIML-----SHGIYSALFLICGVY--DHKT 360

Qy 281 -LHLMQNAVTTIANGTDFC--TSADAFKILSKNSHFTSINCFDPIIFGKVLVVC 337

Db 361 RLKXYSGVVQTMPIFALFPLFTMA-----NISLPTGSFY--GEFVLIG----- 405

Qy 338 FTVFGGLMAFVYNRAFOVAVPVLVAFEA 367

Db 406 -----ARNSNTT-----VAFPA 417

RESULT 14

E69143

hypothetical protein MTH339 - Methanobacterium thermoautotrophicum (strain Delta H)

C/Species: Methanobacterium thermoautotrophicum

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C/Accession: E69143

R/Smith, D.R.; Doucette-Stamm, L.A.; Delughery, C.; Lee, H.; Dubois, J.; Aldredge, T.

J. Bacteriol. 179, 7135-7155, 1997

A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fur

A/Reference number: A69000; MUID:96037514; PMID:93711463

A/Accession: E69143

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-694 <MTH>

A/Cross-References: GB:AE000818; GB:AE000666; NID:G2621384; PIDN:ABB84845.1; PID:G2621

A/Experimental source: strain Delta H

C/Genetics:

A:Gene: MTH339

Query Match

Best Local Similarity 4.8%; Score 110.5; DB 2; Length 694;

Matches 88; Conservative 73; Mismatches 164; Indels 105; Gaps 20;

Db 12 LIALSLAMMFFRTITLLVHIFSLV-----LGLFVCGVLMWLY----- 56

Db 58 VTAALSTFLM---XMTGLINSFPLIMRPSLAPLSAINLLIT--ALWVYHURGE 111

Qy 57 ----DYTNDSIELDERENKCVLGFAIVS--TGITAVLVILFVLRRIKLTVELFQ 109

Db 112 EPLWKKSGNLSLSPVLTVIGSYLMNKSVTVLVLLVF-----IVYITLLE 166

Qy 110 ITNKAISSAPFLFQPLMTFALIFFWLVAVVLSLGTGAAGVMEGGVEKPLSGIR 169

```

Db      167 VFKKIGPA-----TYPVAF-----SISLSL-----LIMNGLPSNY--LIGRD 203
Qy      170 YMSGVHIGL-----IMTSEFLAACQMTIAGAVTTCYFNKSKNDPPHPLISL 219
Db      204 IHMEFYLRKALMAHMDMVGSNAYNACLSTVLEVIYKVLN-----VPEV 252
Qy      220 SILEFYHOGTVKSSFLISVVRIPRIIWMQNALKEQHGALSRYL---PECCYCCFW 275
Db      253 YIFKRY-----GFIGALMPLP--VLLISERILKSDYGFYATLLIFQFSGFIYLGW 303
Qy      276 CLDKYLLHNGNAVTTTANGTDFCTSAKDAFKILSKNS--SHPSTINCGDPIIFLGAV 333
Db      304 C--ROLVALLEPFAAAWVLTGDNRSHKLLFVFMVGYLSHTTAYF---FELVAL 357
Qy      334 LVVCFYVFGIMAFNRYRAF--QVAVPPLLIVAFAYLVA---HSFISVPEYVLDALF 386
Db      358 IPIVVRVWRKLVKVPDSDSGFFNAPSLAVLFFVVVFVFAWYATGAPFKSAVSFFFKTLRSM 417
Qy      387 LCPAVDLETN 396
Db      418 EFSHDMNN 427

```

## RESULT 15

S25942

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - liverwort (Marchantia polymorpha)

N.Alternate names: NADH-ubiquinone oxidoreductase chain 4

C.Species: mitochondrion Marchantia polymorpha

C.Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002

C.Accession: S25942, S31172

R.Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.; Kan

J. Mol. Biol. 223, 1-7, 1992

A.Title: Gene organization deduced from the complete sequence of liverwort Marchantia po

A.Reference number: S25941, MUID:92114051, PMID:1731062

A.Accession: S25942

A.Status: nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-495 <ODA>

A.Cross-references: EMBL:M68929; NID:9786182; PIDN:AC09398.1; PID:9786185

A.Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992.

R.Nozaoto, N.; Oda, K.; Yamato, K.; Ohta, E.; Takemura, M.; Akashi, K.; Fukuzawa, H.; Oh

Mol. Gen. Genet. 237, 343-350, 1993

A.Title: Co-transcriptional expression of mitochondrial genes for subunits of NADH dehydr

A.Reference number: S31171, MUID:93247547, PMID:8483448

A.Accession: S31172

A.Status: nucleic acid sequence not shown

A.Molecule type: DNA

A.Residues: 1-495 <NOZ>

A.Cross-references: EMBL:M68929; NID:9786182; PIDN:AC09398.1; PID:9786185

C.Genetics:

A.Gene: nad4

A.Genome: mitochondrion

A.Introns: 183/2

C.Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C.Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 4.8%; Score 110; DB 1; Length 495;  
 Best Local Similarity 18.9%; Pred. No. 0.25;  
 Matches 78; Conservative 73; Mismatches 152; Indels 110; Gaps 16;

```

Qy      12 ILAALSLAMMFTPTTLLVHFLSVILG-----LFFVGV 51
Db      130 LIAVFCSLDLIFVFFESVLIPIPIIIGWGRORIKAYOFPLYTLGSLFMDLAI 189
Qy      52 WMLVYDYNDLSIELDERENMKCVLGAIVSTGITAVLVLIFFVLRKRITVSEYQRT 111
Db      190 FIFQGTGTDQILLTE-----FSERROILMWIAFFAS 224
Qy      112 NKAISAPFLLFQPLMFAIIFFWVLMVAVLSLGTAG---AAQVMEGGQVEYKPLSG 167
Db      225 SVKVPWVFVHIMLEPAHVEAFAGSVIILKLTGYGFLRPSIMPEPATLYFTP--- 281
Qy      168 IRVMSYHLIGLWTSSEFLAACQMTIAGAVTTCYFNKSKNDPPHPLISLSLIFFYHQ 227

```

```

Db      282 --FIYTLVIAIYTS--LFTIRQIDPKLIA-----YSSVA-----HM 316
Qy      228 GTVYKGSFLISVVRIPRIIWMQNALKEQHGALSRYLFRCCYCCFWLDRX---LHL 284
Db      317 NFVTIGKFSINTQIGESITLML-----SHGLVSSALFLCVGALY---DRKTRIVAK 366
Qy      285 NONAVTTTANGTDFCTSAKDAFKILSKNSHPTSINCGDPIIFLG---KVLVVCFTV 340
Db      367 YGGLVSTMPISSTIF-----LFTLANMSLPQTS-SFIGEFLLVGAFOGNSLVATLAA 419
Qy      341 FGLMANNRYARQVAVPPLLIVAFRA--YLVASFLSVFETVLDALFECFAV 391
Db      420 LGMIL---GAAYSIMLDYNRVFGNFKPNFTLKFSDLMRREV---IFLPLTV 465

```

Search completed: May 27, 2004, 16:37:42  
 Job time : 31 secs